Sea ice segmentation in SAR images using Deep Learning

by

Anmol Sharan Nagi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2020

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see *Statement of Contributions* included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Two publications have resulted from the work presented in the thesis:

1. Anmol Sharan Nagi, Devinder Kumar, Daniel Sola, K. Andrea Scott "RUF: Effective Sea Ice Floe Segmentation using end-to-end RES-UNET-CRF with tailored dual loss" (Submitted to IEEE Transactions on Geoscience and Remote Sensing.)

2. Anmol Sharan Nagi, Manpreet Singh Minhas, Linlin Xu, K. Andrea Scott "A Multi-Scale Technique to Detect Marginal Ice Zones using Convolutional Neural Networks." (Accepted IGARSS 2020.)

Abstract

Sea ice covers over seven percent of the world's oceans. Due to the effect of global warming, Arctic's ice extent has decreased significantly in the past decades. This reduction in sea ice cover is opening new pathways for the international shipping community through the Arctic. Due to the lengthening of the open water season, the Canadian Arctic has also observed a three-fold increase in the shipping traffic in the past few years. Although the ice extent has reduced, the risks and hazards involved in shipping through these regions are still significant. To promote safe and efficient maritime activities in the Canadian Arctic, Canadian Ice Service (CIS) provides information about ice in Canada's navigable waters. CIS uses Synthetic Aperture Radar (SAR) images as one of the prominent sources to gain insights about the ice conditions in Canadian waters. Automated SAR image interpretation is a complex task and requires algorithms to learn complex and rich features. Convolutional neural networks (CNNs) have demonstrated their ability to learn such features and have been used in various image classification, segmentation and object detection tasks. In this thesis, we first propose a method to detect marginal ice zones (MIZs) in SAR images. This method uses transfer learning combined with a multi-scale patch technique to detect the MIZs. The multi-scale patch technique involves generating the segmentation masks over different patch sizes for the same image. These masks are later stacked together and thresholded to generate the final MIZ prediction mask for an image. Later we dive deep into the MIZs and focus on segmenting sea ice floes. We propose a segmentation model optimized for the task of ice floe segmentation in SAR images. The model is based on a fully convolutional architecture with residual connections. In addition to this, a conditional random field is also used as a post-processing step. The whole network is trained end-to-end using a dual loss function. Qualitative and quantitative analysis suggests that our model beats the conventional segmentation architectures for the task of ice floe detection.

Acknowledgements

I would like to thank my supervisor Professor Andrea Scott for the continuous guidance and support that they provided during my master's degree.

I would also like to thank my thesis readers, Professor Alexander Wong and Professor Nima Maftoon for reviewing this thesis and providing their valuable inputs.

I would like to express my gratitude toward my friends Devinder, Kanav, Manpreet, Amanpreet, Apurva, Jasjiv, Sushant, Jashan, Bikram, Manpreet, Hitesh and Harnidh for guiding me though a lot of things, be it personal life or lab-work.

Last but not the least, I would like to thank my mother Sardarni Surinder Kaur and my father Sardar Manmohan Singh to help me in becoming the man I am today. I thank my sisters, Deepiner Kaur, Manjinder Kaur and Jashanjot Kaur for their love and support.

The RADARSAT-2 data used in this paper was provided by the Canadian Ice Service (CIS). All SAR images are copyright to MacDONALD, Dettwiler and Associates Ltd. (2010)—All Rights Reserved. RADARSAT is an official mark of the Canadian Space Agency. I also wish to thank the Canadian ArcticNet program for funding this research.

Dedication

This is dedicated to my parents.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Sea ice is one of the greatest physical constraints for shipping activities in the Arctic. Due to lengthening of the open-water season, maritime traffic in the Arctic has increased three-fold over the past few years [1]. The story is similar for the Canadian Arctic as well, with Hudson Strait being the most traffic prone area [3]. Despite the reduction in extent and thickness of ice across the Arctic, the risks and hazards involved in Arctic shipping remain significant.

The transition region from ocean water to sea ice is called the marginal ice zone (MIZ). MIZs are defined as regions with sea ice concentration between 15% to 80% [65]. MIZs are very dynamic regions containing ice floes, ice eddies and ice drift etc. As the ice melts, the ice pack becomes more mobile, allowing hazards such as ice floes to break away. These ice floes move at high speeds and can cause damage to vessels and man-made structures [8]. Ice charting is done by the Canadian Ice Services (CIS) for estimating ice concentration to make management decisions for ensuring safety and efficient maritime activities in the Canadian Arctic [93]. CIS uses data from various data sources for the production of ice charts and development of guidelines for mariners. One of the prominent data sources used for the task of sea ice monitoring is Synthetic Aperture Radar (SAR) which provides high spatial resolution images irrespective of the daylight conditions [99]. Due to the recent advancements in satellite technology and launch of new SAR satellites e.g. RCM (RADARSAT Constellation Mission), the volume of available SAR data is increasing. For the sea ice community, these data are typically interpreted by experts.

To process the vast volumes of available data in real-time, there is a need for automated methods for sea ice detection. Previous studies by Steer *et al.* [101], Toyota *et al.* [109], Holt *et al.* [41], Hwang *et al.* [46], Clausi *et al.* [18] and Zhu *et al.* [127] used image

processing techniques coupled with machine learning on SAR images for the task of ice-water segmentation. SAR images usually have an intrinsic speckle noise in them due to the coherent nature of imaging process [20]. The presence of this speckle noise is one of the main factors limiting the accuracy of the previously proposed methods [101, 127]. To circumvent this issue, instead of using SAR images some studies (Hall *et al.* [32], Lu *et al.* [67], Heyn *et al.* [38, 39] and Wang *et al.* [115]) used data from vessel mounted cameras for identifying ice floes. This approach does not have the issue of speckle noise contamination. However, geometric error compensation is required to tackle the problem of underestimation/overestimation of ice cover caused due to oblique sensor placement [67].

Recently, the use of data driven approaches such as neural networks has gained traction in the image processing community due to availability of large image datasets and advancements in graphics processors. Neural networks have shown to outperform traditional machine learning techniques for various image processing tasks such as classification, segmentation and object detection etc. [5]. Specific to the task of remote sensing, neural network models have been applied for image classification, object detection, segmentation, super-resolution and de-noising etc.

This research aims to develop automated methods to ease the tedious task of manually interpreting SAR images for sea ice detection. In this thesis two approaches for sea ice detection have been explored. We annotated a dataset of over 150 SAR images for the task of MIZ and sea ice floe detection. The first approach focuses on detecting MIZs using a convolutional neural network (CNN). Here, the convolutional neural network is used as a fixed feature extractor (Chapter 3). Later, we propose a fully convolutional network optimized for the task of ice floe segmentation. We performed various tests to demonstrate the performance of our methods.

1.1 Thesis Contribution

The main contributions of this thesis are as follows:

- Network based transfer learning for marginal ice zone (MIZ) detection in SAR images using convolutional neural networks (CNNs) was explored. Various classification models were investigated for the above task. A mutliscale technique is proposed to generate the MIZ segmentation masks. Results from visual inspection suggest that the presented approach was successfully able to detect marginal ice zones .
- RUF (Residual-UNET-CRF): a tailored convolutional neural network for sea ice floe segmentation in SAR images is presented. RUF features an architecture inspired from

residual connections [37], UNETs [87] and conditional random fields [106]. The presented method is trained using a dual loss function. In comparison to FCN8 [66] and DeepLabV3 [14], RUF achieved an impressive average improvement in performance with an F1 score of 74.52 (108.5% increment), mIOU of 66.26 (120.5% increment) and mPA of 77.88 (110.3% increment).

1.2 Thesis Outline

The remainder of thesis is organized as follows: In Chapter 2 background related to sea ice, remote sensing, convolutional neural networks, various classification and segmentation models, and transfer learning is discussed. Next, the network based transfer learning approach for marginal ice zone detection is presented in Chapter 3. Later in Chapter 4, RUF architecture for ice floe segmentation is explained. Finally, in Chapter 5, summary and conclusion of this thesis are presented.

Chapter 2

Background

This chapter is divided into two sections, sea ice remote sensing and convolutional neural networks (CNNs). The first section provides the background regarding sea ice, marginal ice zones (MIZs) and, synthetic aperture radar (SAR). The second section provides the background of convolutional neural networks along with various classification and segmentation architectures used in this study. In the last section various transfer learning techniques are discussed.

2.1 Sea Ice and Remote Sensing

2.1.1 Sea Ice

Sea ice occupies about 7 percent area of the world's oceans [112]. Even though sea ice is a local phenomenon occurring near the polar regions of the earth, it has a prominent impact on global climate, flora, and fauna. Moreover, sea ice impacts the global economies as it has a significant impact on polar navigation and resource extraction.

The formation of sea ice is a thermodynamic process. The freezing of sea-water leads to the formation of tiny crystals called Fazil. Most salt and other ions in the seawater drain out upon crystallization. As ice floats on water, the salinity of water under the sea ice cover becomes higher than normal, and this water freezes at lower temperatures due to dissolved impurities [72, 107]. This layer of cold water creates a barrier between the warmer ocean water and the newly formed sea ice, allowing the ice sheet to grow from below and increase in thickness [81]. When sea ice grows, salt and other ions which were

Figure 2.1: An image of Arctic sea ice. (Photo by Paul Nicklen, National Geographic Creative [76])

earlier expelled during the formation of the ice layer get trapped in the ice [107]. These ions crystallize and concentrate together to form brine droplets that give sea ice its opaque and porous texture [107].

When ice floats on the ocean it becomes a part of the oceanographic system, which means that it is in constant motion. Convergent and divergent forces act on sea ice due to winds and currents. Divergent forces release ice pressure and lead to thinning of ice while convergent forces increase the thickness via ridging as illustrated in Fig. 2.2. These dynamic forces lead to pressured ice, the formation of ridges and breakup-up of icepack into floes; posing hazards for vessels and offshore structures [108, 60].

2.1.2 Marginal Ice Zones (MIZs)

Sea ice concentration is defined as the coverage of sea ice in a given geographic area, and it ranges from 0 percent to 100 percent. The transition regions from consolidated ice to open water areas are called Marginal Ice Zones.

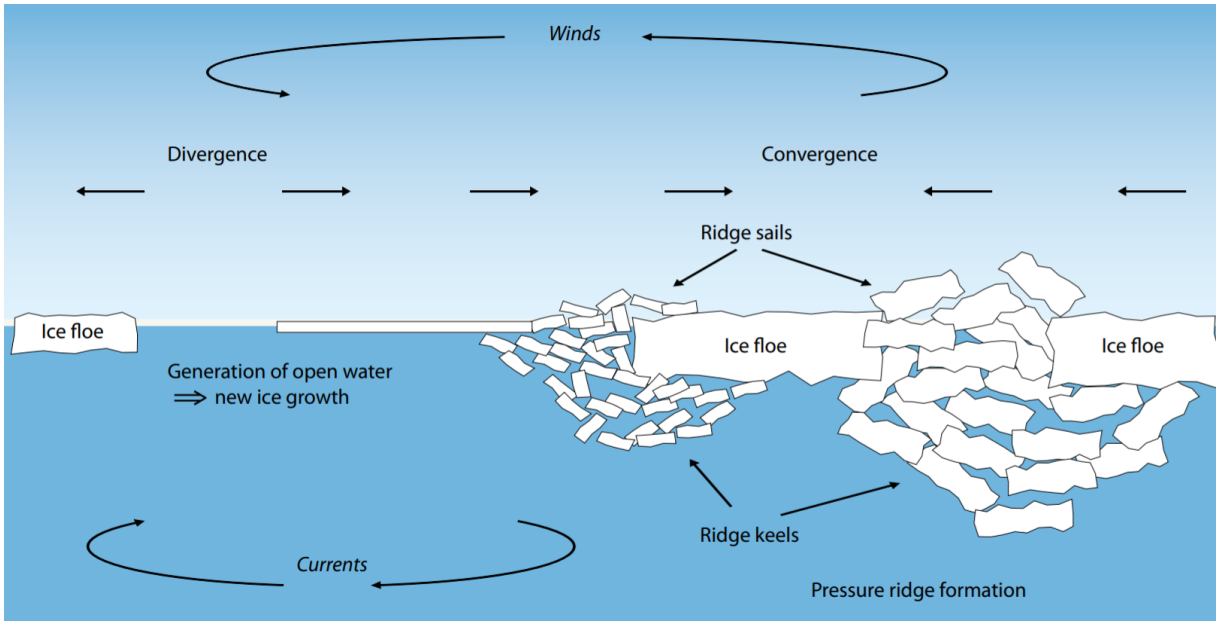


Figure 2.2: Illustration of different processes contributing changes in sea ice thickness. Reproduced from AMAP, 2012. [73]

MIZs are regions of intermediate ice concentration (eg. between 15% and 80%) [65]. In these regions, there are interactions between air, ice, and ocean. The MIZ cover is very sensitive to the wind. When the wind is off the ice cover, strips of ice bands drift and diffuse away from the main ice cover. However, the ice edge is very sharp when the wind is towards ice cover. MIZs have large seasonal and year-to-year variations in cover. In summers, waves can penetrate deep into the ice cover breaking it into ice floes. These ice floes are discrete masses of frozen seawater. Their lateral sizes span orders of magnitude, from small ice floes of the order of centimeters to fused ice floes with size on the order of kilometers. The ice floes are known to cause ice jams and damage the hulls of ships. Detection of ice floes is important for managing offshore Arctic operations such as:

- Initializing ice-ocean models and validation of generated results [23, 70, 91, 92].
- Detection of hazards for maritime traffic [78, 69].
- Provide an early warning for a case of ice compaction event [116, 78].
- Conduct empirical analysis for estimation of ice load on stationary structures [55, 78].

2.1.3 Synthetic Aperture Radar (SAR)

Synthetic aperture radar (SAR) is a form of radar that uses 'active' remote sensing technology in the microwave region (low frequency e.g. 5 GHz) to capture images. SAR sends electromagnetic radiation pulses to the ground and collects the back-scatter as measurements. Due to the active nature of measurements, SAR does not require solar illumination for scene capture around the globe and can be used irrespective of daylight conditions. SAR systems have all-weather capability as they use longer wavelengths than optical systems, allowing them to penetrate clouds and haze to acquire images [99]. To capture images of large areas at high resolutions, SAR systems use the motion of radar antenna along with signal processing to 'synthesize' the radar's aperture allowing it to create high-resolution images with relatively small antennas. These characteristics make SAR systems an optimum candidate for remote sensing operations in the Arctic, where clear skies are scarce.

SAR operates at swath widths between 18 and 500 km, and spatial resolutions between 1 and 1000 m. This property allows them to be used for observing both regional and local variations in the sea ice cover parameters. Finally, as SAR systems measure back-scatter from the transmitted radar pulses, this allows data collection in different polarization modes. Polarisation is the orientation of electromagnetic radiation transmitted and received by the antenna. SAR systems have four polarisation configurations:

- **HH**. Transmission: Horizontal, Reception: Horizontal
- **HV**. Transmission: Horizontal, Reception: Vertical
- **VH**. Transmission: Vertical, Reception: Horizontal
- **VV**. Transmission: Vertical, Reception: Vertical

HH and VV are called like-polarised, as transmission and reception polarisation is the same. HV and VH are called cross-polarized, transmission, and reception polarisation are different. SAR systems like RADARSAT-2 can coherently transmit and receive both horizontal and vertical polarization. Different polarization aid in gathering different but correlated information about the same surface. Scattering of radiation from a given surface depends upon the polarisation of the incident radiation. Using multiple polarization signatures may help in selecting the optimum polarization for imaging a particular feature and to differentiate between different features of interest.

Canadian Ice Service (CIS) utilized data obtained from RADARSAT-2 satellite to monitor the environment, manage natural resources and perform coast line surveillance from

April 2013 till June 2020. Some relevant specifications of the RADARSAT-2 satellite are given below:

- **High spatial resolution**: 1 m - 100 m (depending on operation mode)
- **Swath width**: 18 km - 500 km (depending on operation mode)
- **Polarisation**: Single, Dual or Quad
- **Sensor specifications**: C-band wavelength 56 mm, frequency 5.4 GHz

2.1.4 Sea ice and SAR imagery

SAR systems image a surface by measuring the back-scatter off the medium. This back-scatter comprises of surface scattering and volume scattering. Surface scattering occurs on the surface boundary between two mediums, while volume scattering occurs upon penetration into the medium. In the case of mediums with high ion concentration (ocean water, young sea ice, wet snow), the dielectric constant is high. Materials with a higher dielectric constant exhibit a greater attenuation to the penetrating electromagnetic fields, thus very low volume scattering. The dielectric constant is a dimensionless complex quantity with real and imaginary parts. The amount of surface scattering depends upon the roughness of the surface; such that, rougher surfaces have a stronger back-scatter.

For sea ice observations, the main frequency bands covered by SAR are L-band (wavelength 150 - 300 mm, frequency 1 - 2 GHz, C-band (38 - 75 mm, 4 - 8 GHz), X-band (24 - 38 cm, 8 - 12.5 GHz) and Ku-band (17 - 24 mm, 12.5 - 18 GHz). The appearance of sea ice in SAR images is affected by the radar polarization, viewing geometry, frequency, of the ice type, noise intensity, and spatial resolution [111, 4, 25]:

- **Polarization**: Ocean clutter is better suppressed in HH polarization than VV polarization. Hence, HH polarization is preferred for differentiating between ice and water. During the presence of high wind speeds, ocean water can appear brighter than sea ice in like-polarization. In such cases, cross-polarization can be used [56, 25].
- **Incidence angle**: As incidence angle increases the contrast between the appearance of rough and smooth ice increases. HH polarization is strongly effected by the incidence angle than VV polarization [51].

- **Frequency**: Penetration depth of L-band is up to the depth of tens of centimeters, while that of C-band and X-band is of the order of centimeters [33, 27, 10]. X-band is more sensitive to the surface and sub-surface layer properties than C-band due to lower penetration. L-band is able to better differentiate between structures such as brash ice, ridges, and rubble than C-band.

- **Noise intensity**: Minimum detectable radar intensity determines the noise intensity. In cross-polarization modes, the back-scatter intensity of both smooth young sea ice and open water can fall below the noise intensity. Thus, making it difficult to differentiate between early stages of ice growth [25].

- **Spatial resolution**: To obtain detailed information about the ice surface, spatial resolution better than 10 m is required. Data products of higher spatial resolutions have narrower swaths. But, wider swaths are preferred for operational mapping along shipping routes, with the drawback that smaller ice structures cannot be recognized in the coarser spatial resolutions [25].

2.1.5 Study Area

The geographical area under study is comprised of Hudson Strait and Ungava Bay located in Eastern Canada, and their outflow into the Labrador Sea in the North Atlantic. A picture of this area is shown in Fig.2.3. Hudson Strait links the Hudson Bay with the Labrador Sea and the Atlantic Ocean. Hudson Strait serves as the passage between northern seaports of Ontario and Manitoba with the Atlantic Ocean. Hudson Strait on average is 150 km wide with a minimum width of 70 km at the eastern end [26]. The strait a steep-sided U-shaped profile and ranges from 300 to 900 m in depth. Hudson Strait has a two-way circulation system. Seawater from the Labrador Sea runs from east to west along the northern half of the strait into the Hudson Bay, while fresher water with lower salt content leaves the strait from the southern half [103, 102]. The net volume of flow (the difference between outflow and inflow) through the Hudson Strait is about 42000 cubic meters per second into the Atlantic Ocean.

Hudson Strait's ice cover lasts for about 8 months of the year. A typical ice cycle for the formation and melting of sea ice in this area is as follows: The freeze-up usually progresses from the western end of the Hudson Strait towards the eastern end. First sea ice intrusions can be observed in later October and early November through the Foxe Channel. In situ freeze-up is observed in mid-Novermber with the formation of land-fast ice along the coasts of Baffin Island, Quebec, and Ungava Bay [42, 95]. By early December most of the Strait

Figure 2.3: Geographic area of study including Hudson Bay, Hudson Strait, Ungava Bay and Labrador Sea. Red polygon represents the geographic extent of a SAR scene described below.

is covered with drift ice, except the land-fast ice in the Ungava Bay inlets and Baffin Island coasts. The thickness of land-fast ice can reach a maximum of 110 cm to 160 cm [95]. The area witnesses a slow melting of the ice pack in early July, which increases rapidly though summer months [90]. There is a complete meltdown of sea ice usually during the second week of August, with the availability of vast open waters [12].

A sample scene from the dataset is shown in Fig. 2.4. This scene captures the eastern end of the Hudson Strait and the tip of Labrador. The scene contains different ice types, water, and land. It can be observed that the smooth surface of the ocean appears very dark while the rough land and ice appear quite bright. As HH scenes are strongly affected by the incidence angle, water in the scene appears dark in the far left while it appears bright in the far right. As the scene is captured in late May, a thick ice cover with embedded

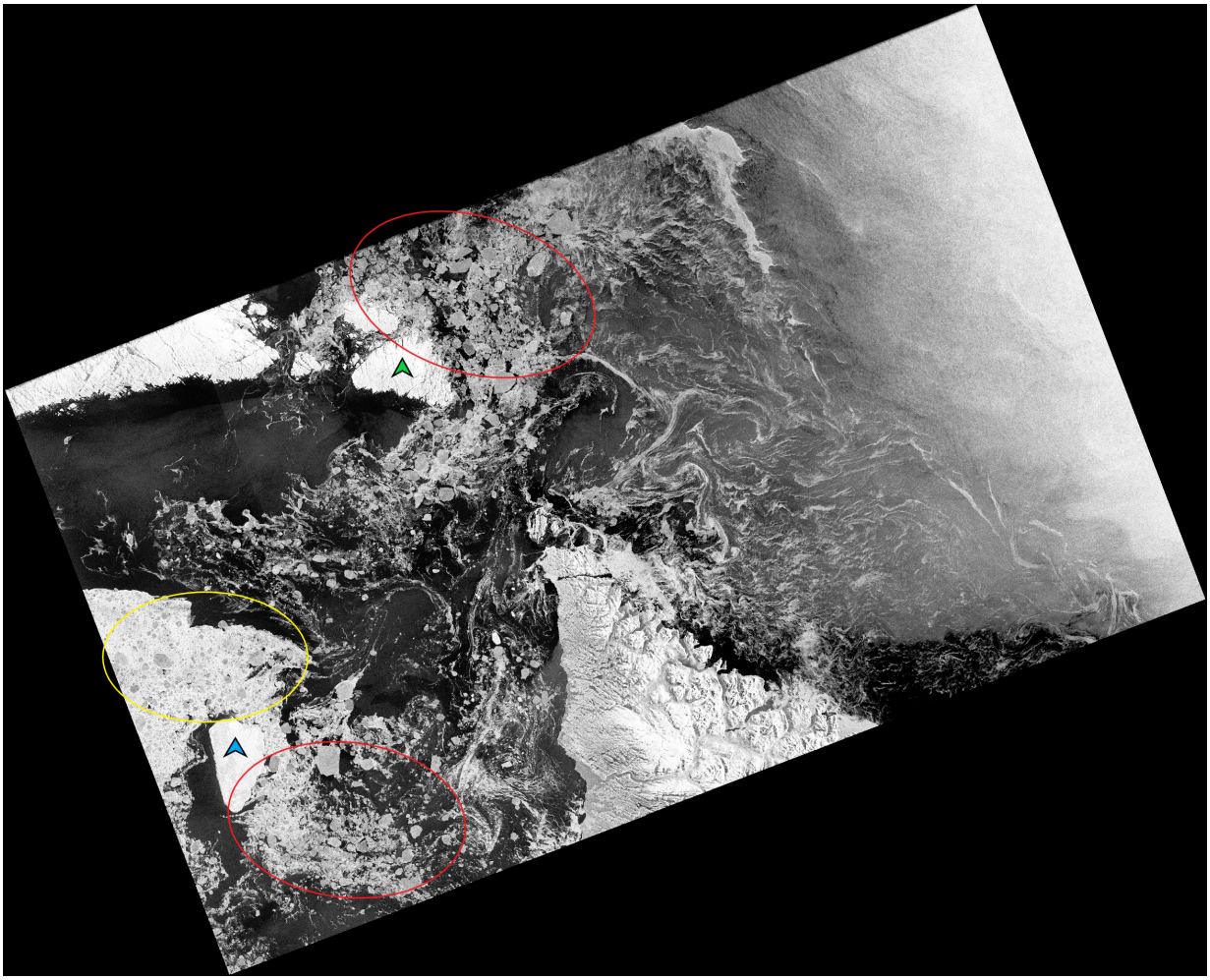


Figure 2.4: A sample RADARSAT-2 SAR scene captured on May 26, 2011 in HH polarisation. Akpatok island (blue) and Resolution island (green) are highlighted for reference. Some floes embedded in ice cover are marked with a yellow oval. Some floes with homogeneous boundaries are marked with red ovals.

floes (marked with yellow oval) can be observed next to the Akpatok island. Marginal ice zones containing ice eddies and ice floes can be seen throughout the image. But it is very difficult to discern between different floes (marked with red ovals next to Akpatok island and Resolution island). It can be noticed that the ice floes show a lot of variation both in their shapes and sizes. Due to the complex nature of such scenes, the use of intelligent models for detecting various features in SAR images arises.

2.2 Convolutional Neural Networks

Convolutional neural networks (CNNs) are a type of neural network which are used to process data that has a grid-like topology [30]. Time-series data can be thought to be composed of 1-D grids, while images being 2-D grids and videos as 3-D grids. Currently, CNNs are one of the most popular image recognition architectures. CNN architectures were initially inspired from understanding of the cat visual cortex [45]. In the visual cortex, each neuron is sensitive to a specific region in the visual field. The organization of neurons is in a layered architecture, which leads to a belief that different layers are used to understand different levels of abstraction in the images. This principle is not much different from machine learning's hierarchical feature extraction. A typical CNN architecture comprises of convolutional layer, pooling layer, activation layer, batch-normalization, dropout, and fully connected layer.

2.2.1 Convolutional Layer

The name 'convolutional' come from the utilization of the mathematical convolution operation (a linear operation/activation). The discrete time convolution of two 1-D discrete time signals f and g is defined as follow:

$$c[n] = f[n] * g[n] = \sum_{v=-\infty}^{\infty} f[v]g[n-v], \tag{2.1}$$

such that n & v are variables with discrete values. In case of a 2-D convolution the above operation becomes:

$$c[m,n] = f[m,n] * g[m,n] = \sum_{v=-\infty}^{\infty} \sum_{u=-\infty}^{\infty} f[v,u]g[m-v,n-u], \tag{2.2}$$

such that m, n, u &, v are variables with discrete values. In case of an image, these values would range between the image dimensions and not $(-\infty, \infty)$. A visual example of 2-D convolution is shown in Fig. 2.5.

2.2.2 Activation Function Layer

Convolutional layer linearly maps the input to output. As most of the real world problems involve some degree of non-linearity, a non-linear property must be induced in the neural

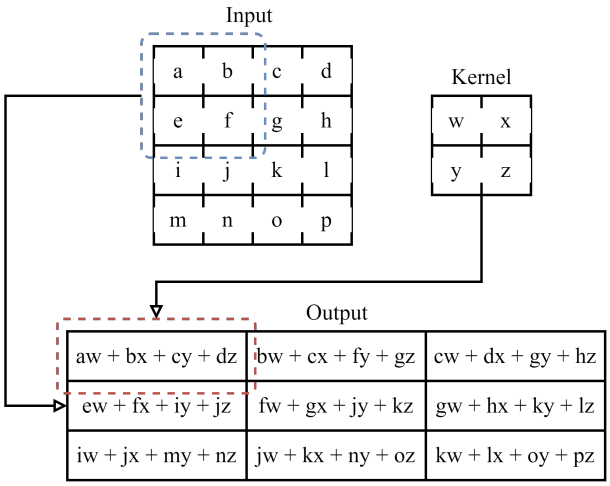


Figure 2.5: An example of 2-D convolution.

networks. Activation functions are responsible for inducing this non-linearity. Activation functions are an abstraction of action potential in a neuron; such that these are responsible in deciding if a neuron should fire or not. In case of neural networks, the activation function decides if the information received from the neuron in previous layer is relevant to the task or not. The activation function φ is applied element-wise to input x with weight W and bias b. The output C becomes:

$$C = \varphi(W * x + b). \tag{2.3}$$

Currently, a default recommendation for the activation function is the rectified linear unit or ReLu [74, 50]. In the past, neural network architectures have used sigmoid and tanh as activation functions. ReLu activation function can be defined as:

$$\varphi(z) = max(0, z). \tag{2.4}$$

ReLu is simply a max operator, such that if the input z is greater than 0, then the information is useful, else not. ReLu is a piecewise linear function with two linear pieces and allows the model to be easily optimized using gradient-based methods. The output of ReLu is not squished between any bounds as in the case of sigmoid and tanh, avoiding the vanishing gradient problem [30]. A visual representation of ReLu activation function is shown in Fig. 2.6.

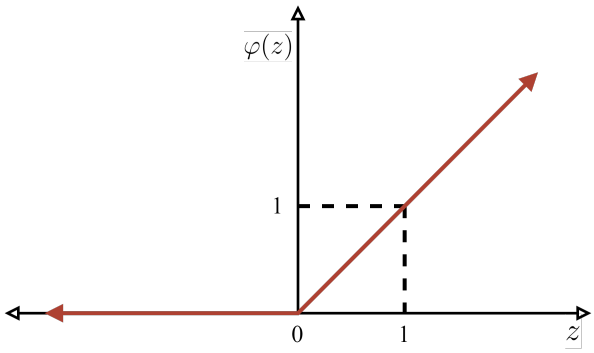


Figure 2.6: ReLu non linearity visualized.

2.2.3 Pooling Layer

Pooling operation is applied to the feature maps obtained from the non-linear activation layer. Pooling is a spatial reduction operation that produces another layer with the same depth of that of the original layer [126]. The values in the new layer represent the statistical summary of the nearby regions in the original layer.

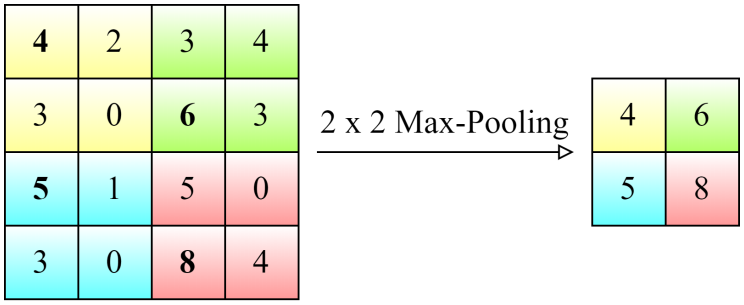


Figure 2.7: Illustration of max pooling operation on a 4×4 layer with a region/kernel size of 2×2.

As the summary of the statistics of a local region ($k \times k$ pixels) in the previous layer is placed 1 pixel apart in the new layer, the computational efficiency of the next layer is improved by roughly k times [30]. Pooling is applied to the input image or feature map by dividing it into non-overlapping patches and then sampling each patch. A visual example of max pooling operation is shown in Fig. 2.7.

2.2.4 Batch Normalization (BNorm) Layer

Batch normalization (BNorm) [48] is a type of adaptive reparameterization, devised to improve the training of models with many layers. BNorm addresses the problem of vanishing and exploding gradient where the activation gradients of the consecutive layers either increase or decrease in magnitude.

Considering the output of previous layer v as a batch with m examples, such that $v^{(1)}, v^{(2)}, v^{(3)}, ...v^{(m)}$ are the m activation values. We calculate the mean(μ) and standard deviation(σ^2) for the given batch and normalize the input v as:

$$\mu \longleftarrow \frac{1}{m}\sum_{i=1}^{m} v^{(i)}, \tag{2.5}$$

$$\sigma^2 \longleftarrow \frac{1}{m}\sum_{i=1}^{m}(v^{(i)} - \mu)^2, \tag{2.6}$$

$$\hat{v}^{(i)} \longleftarrow \frac{v^{(i)} - \mu}{\sqrt{\sigma^2 + \epsilon}}.$$

where ϵ is added to regularize the cases with same activations. For better optimization, two trainable parameters, a shift(α) and scale(β) are applied to $\hat{v}^{(i)}$ [48]:

$$a^{(m)} \longleftarrow \alpha\hat{v}^{(i)} + \beta. \tag{2.7}$$

2.2.5 Dropout

Dropout [100] uses probability to drop certain neurons in a network while training. A node is dropped from the network with probability p, where p is a parameter. These nodes are dropped temporarily. Dropout is used as a regularizer in the network. It can be thought of as an ensemble method as dropping various neurons creates different networks.

2.2.6 Fully Connected Layer

Multiple fully connected layers can be connected to the final spatial layer such that each neuron in the previous layer is connected to the next layer. Each fully connected layer consists of matrix multiplication of input from previous layer x_{l-1}, similar to equation 2.3:

$$C_l = \phi_l(W_l x_{l-1} + b_l), \tag{2.8}$$

where, C_l is the output, W_l are the weights and b_l is the bias of layer l in the network.

Since all the neurons in both of the layers are densely connected, the number of trainable parameters increases rapidly when fully connected layers are used in a network.

2.2.7 CNN Architecture

A typical CNN architecture combines the components explained above; various convolution and relu layers are stacked together, followed by a pooling layer. The last layer of the network is application-specific. In case the task is classification, the last layer is composed of a fully connected layer. While for the task of segmentation, the last layer is a convolutional layer, resulting in a fully convolutional network. Batch Normalization and Dropout layers are also incorporated in the architecture as explained above to enhance model training and generalization abilities.

2.3 Classificaiton

Classification in terms of computer vision refers to the process of categorizing various input images to given classes. In this research, CNN based image classification networks, namely VGG, ResNet, and DenseNet have been used. Brief details about each network are provided below:

2.3.1 VGG

VGG was one of the top-performing networks for the ISLVRC 2014 competition and was developed by Simonyan and Zisserman [98]. VGG focuses on reducing the convolutional kernel sizes but increasing the network depth with more layers. VGG architecture uses a stacked configuration of 3×3 kernel to simulate the receptive field of 5×5 and 7×7 receptive fields. A deeper network has more nonlinearity as compared to a shallower network due to the presence of a greater number of ReLu layers [2].

For comparison, three stacked convolutional layers with kernels of size 3×3 require $3(3^2C^2) = 27C^2$ parameters, in place of one convolutional layer with kernel size 7×7, that requires $7^2C^2 = 49C^2$ parameters. Therefore, the number of parameters are reduced drastically and can be put to use in increasing the model depth.

The problem with a greater number of layers is that, as the number of layers increases, the sensitivity of parameter weight initialization also increases. This may cause instability in the model and impede training. To tackle this issue VGG trains a subset of layers first, before training the model in the deeper configuration. Refer to [98] for more details.

2.3.2 ResNet

ResNet was proposed by He *et al.* [37]. In most of the feedforward network the lth layer is connected to $(l+1)th$ layer, but ResNet architecture connects lth layer and $(l+r)th$ layer where $r > 1$. Fig. 2.8 shows a residual block with this behavior. These skip connections allow copying between layers which enables efficient gradient propagation through the network. Zagoruyko *et al.* [121] supports increasing the number of filters in each layer rather than increasing the number of layers to improve the performance. Refer to [37] for more information.

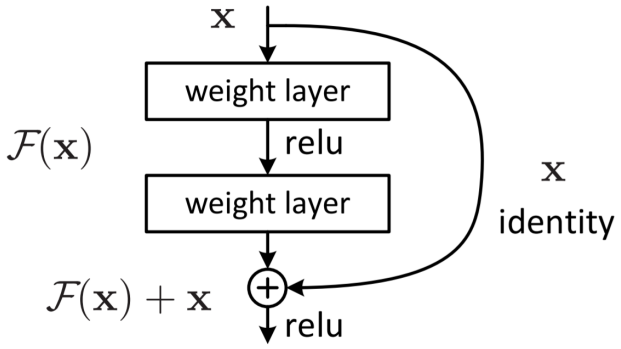


Figure 2.8: Residual learning block. Reproduced from [37].

2.3.3 DenseNet

DenseNet [44] concatenates all the feature maps of previous $(l-1)$ layers to the lth layer. This architecture is similar to ResNet as information from previous layers is used via skip connections to extract different levels of abstraction. Instead of adding the output from the previous layer in case of ResNet, DenseNet concatenates the output from previous layers Refer to [44] for more information.

2.4 Segmentation

Segmentation in terms of computer vision refers to the process of performing pixel-level labeling about given object categories for all the pixels in an image. In this research, we focus on semantic segmentation which is the task of clustering various parts of an image together that belong to the same object category. UNET and Conditional Random Fields (CRFs) architectures used in this thesis are discussed in the next section.

2.4.1 UNET

UNET was proposed by Ronneberger *et al.* [87] for segmenting biomedical images. UNET architecture comprises of two parts, an encoder (downsampling path) and a decoder (upsampling path). UNET is a fully convolutional architecture, the encoding path extracts features using 3×3 convolutions, reducing the dimensions for feature maps. The decoding path increases the dimensions of feature maps while reducing the feature maps. Similar to DenseNet, UNET uses skip connections for transferring information across various encoding-decoding blocks. Skip connections help in integrating the location information in the downsampling paths to the contextual information in the upsampling paths. Refer to [87] for more information.

2.4.2 Fully-Connected Conditional Random Fields: FC-CRFs

In a semantic segmentation task, the pixel-wise predictions of the CNN models are prone to having inaccurate boundaries. To reduce inaccuracies in the boundary, global and contextual information models such as CRFs can be used in conjunction with the CNNs [14].

In the case of semantic segmentation the label of each pixel is $x_i \in \{1, ..., X\}$, where i is a pixel in image I with N pixels. In a fully connected CRF, as one pixel in the image is connected with all other pixels, the model energy function can be written as

$$E(\hat{x}|I) = \sum_{i=1}^{N} \Psi_i(\hat{x_i}|I) + \sum_{i=1}^{N} \sum_{j=1}^{N} \Psi_{i,j}(\hat{x_i}, \hat{x_j}|I). \tag{2.9}$$

The unary potential $\Psi_i(\hat{x_i}|I) = -\log P(x_i)$, where $P(x_i)$ is the probability of label assignment at pixel i as computed by the CNN. The pairwise potential accounts for joint

distribution of pixel pairs i, j as

$$\Psi_{i,j}(\hat{x_i}, \hat{x_j}|I) = \mu(x_i, x_j) \left[w_\alpha \exp\left(-\frac{|\mathbf{p}_i - \mathbf{p}_j|^2}{2\theta_1^2} - \frac{|\mathbf{c}_i - \mathbf{c}_j|^2}{2\theta_2^2} \right) + w_\beta \exp\left(-\frac{|\mathbf{p}_i - \mathbf{p}_j|^2}{2\theta_3^2} \right), \right. \tag{2.10}$$

where $\mu(x_i, x_j)$ is a compatibility transformation such as in Potts model: $\mu(x_i, x_j) = 1$ if $x_i \neq x_j$, and zero otherwise. The expression contains two Gaussian kernels. The bilateral kernel depends on both pixel positions ($\mathbf{p}$) and RGB colour ($\mathbf{c}$), whereas the unilateral kernel depends only on pixel positions. Bilateral kernel forces the pixels within close proximity and similar color to have similar labels, while the unilateral kernel emphasizes on edge smoothness. $w_\alpha, w_\beta, \theta_1, \theta_2$ and θ_3 are the learnable parameters in the model. Refer to [57] for more information.

2.4.3 Convolutional Conditional Random Fields: Conv-CRFs

Conv-CRFs have an add-on assumption of conditional independence over the FC-CRFs. In the Conv-CRF model, two pixels i, j are considered conditionally independent when their L1 norm is greater than a threshold: $d(i, j) > k$ where $d(\cdot)$ is the L1 norm and k is the distance threshold or the filter size. It means that the pairwise potential is zero for all the pixels which are at a distance greater than the threshold k. Hence, energy function in equation 2.9 can be rewritten as

$$E(\hat{x}|I) = \sum_{i=1}^{N} \Psi_i(\hat{x_i}|I) + \sum_{i=1}^{N} \sum_{j=1}^{k \times k} \Psi_{i,j}(\hat{x_i}, \hat{x_j}|I). \tag{2.11}$$

This greatly reduces the computational complexity of the pairwise potential. Teichmann *et al.* also implemented an efficient message passing framework similar to 2d convolutions of CNNs that can be efficiently implemented using convolutional libraries [106].

The capability of Conv-CRFs to learn Gaussian features is demonstrated by [106], where the input features of the unilateral/smoothness kernel are initialized same as hand-crafted features [59], but are later learned and modified during the CRF training.

2.5 Using Pre-trained Models

Supervised training of deep learning models requires lots of labelled data. Image labelling is a demanding task and requires personnel armed with domain specific knowledge (similar

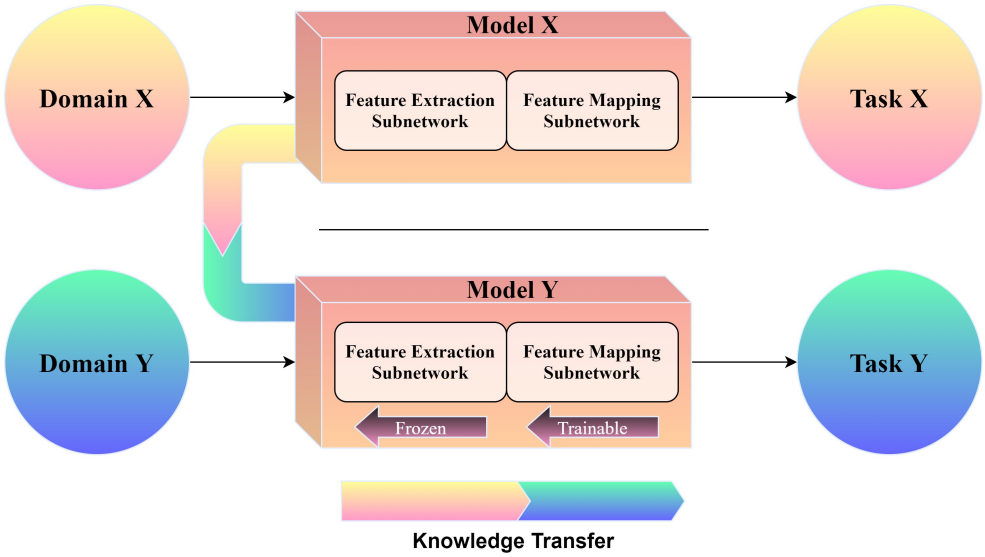


Figure 2.9: Illustration of network based transfer learning from Domain X to Domain Y. Model X is pre-trained for Task X. Model Y is build over Model X by freezing some of the initial layers and training the later layers by back-propagation.

to the issue faced by CIS for ice charting). For most of the practical problems, it is rare to find a labelled dataset of sufficient size to train a deep neural network of choice. A popular technique to solve this issue is by the use of transfer learning. A visual representation of learning using pretrained models is provided in Fig. 2.9.

Transfer learning is a machine learning technique that aims on applying the knowledge gained by solving one problem to solve a different but related problem. It has been observed that the earlier layers of a deep CNN learning primitive features such as Gabor features, color segments and edges [120, 2], while the deeper layers learn more complex features related the specific task. Deeper features (specific to dataset) are much more sensitive to task at hand than the primitive features (common across various image datasets), which are learned in almost all tasks [120, 2]. This knowledge allows us to use a pre-trained model by fixing the learned weights for some of it's initial layers and training only the weights of the deeper layers for our dataset. The two major transfer learning scenarios are as follows:

- **Fixed-Feature Extraction**: When a CNN is used as a fixed feature extractor, the last layer is replaced by another task-specific layer. Assuming total n layers in the network, weights of the initial $n-1$ layers will be kept frozen (they will not be updated) during training. While, only the newly added last layer or the softmax

classifier is trained on the available datatset.

- **Fine-Tuning**: In this scenario, weights of the last n layers (which were frozen otherwise) are updated during training. The number of layers of be trained depends upon the size of dataset and the number of trainable parameters in the layers. Fine-Tuning is prone to overfitting when an excessive number of layers are trained on a relatively smaller dataset. To avoid this issue, the learning rate of the network is kept quite small.

2.6 Sea Ice and Deep Learning

Sea ice charting is typically done by national ice services to identify the boundaries between ice and open water, and to identify the dominant ice types an ice concentration for a given region. In the past years, due to the improvement in both aerial and remote sensing sensors, numerous studies using sea ice data have emerged [34]. These studies cover ice-water segmentation [63, 40, 117], ice concentration estimation [113, 16, 52], ice thickness estimation [97, 53] and ice type classification [122, 79] in detail.

For the task of ice floe detection, various researchers have used different data platforms. Studies conducted by Hall *et al.* [32], Lu *et al.* [67], Heyn *et al.* [38, 39] and Wang *et al.* [115] used vessel mounted camera sensors to obtain photographic data to identify ice floes. However, due to the oblique sensor placement, accurate measurement of sensor height, tilt, and focal length are required to calculate the geometric distortion. Moreover, compensation to ship sway caused due to possible floe collisions and far range shadowing is required for the success of these methods. Hence, later studies used SAR images to avoid such problems.

Images obtained from SAR provide a continuous stream of high spatial resolution data irrespective of the weather conditions and natural illumination. Earlier studier [101, 109, 41, 46, 18, 127] aimed to solve the problem of ice floe identification in two steps. The first step involved the ice-water segmentation while the second step involved delineating different floes. Studies by Steer *et al.*and Toyota *et al.* [101, 109] involved different thresholding methods for sea ice segmentation followed by morphological dilation/erosion operations to split different floes. Holt *et al.* [41] used local dynamic thresholding [35] and shrinking/growing algorithm [110] for floe segmentation. Hwang *et al.* [46] proposed a segmentation technique using Kernel Graph Cuts (KGC) [88] for ice-water segmentation and a combination of distance transformation, watershed [54] and a rule-based boundary re-validation processing for floe splitting [85]. Graphical models such as Markov Random

Field and Conditional Random Field have also been used for the task of sea ice segmentation [18, 127]. Due to the presence of speckle noise in SAR images, it is quite difficult to use traditional machine learning techniques. Depending on sea ice conditions, incidence angle, and sea surface roughness, SAR image distributions can vary from Gaussian to Gamma. CNNs have proven to be good at learning the low and high-level abstract features from raw images. Recently, Singh *et al.* compared various segmentation models (e.g. DeepLab [14], UNet [87], SegNet [6], DenseNet [46]) for the task of river ice floe segmentation. Zhang *et al.* [123] introduced a convolutional network with dual attention streams for ice segmentation in rivers. Research from both Singh *et al.* and Zhang *et al.* focused on segmenting ice floes in rivers using optical data collected from drones.

The main objective of this thesis was to develop automated methods for sea ice segmentation in SAR images. In this research we use SAR data for the task of MIZ detection and ice floe segmentation. We developed two approaches using CNNs to solve this task. In the first approach, we developed a segmentation pipeline using a multi-scale patch technique to detect MIZs in SAR images. In the second approach, we developed a new segmentation architecture to detect individual ice floes in the MIZs of SAR images.

Chapter 3

MIZ detection using transfer learning

Shipping traffic has grown steadily in the Arctic in recent years. One of the reasons for this increased traffic is the lengthening of open water season, which is accompanied by increases in the area covered by intermediate ice concentrations, or marginal ice zones (MIZs). In this chapter, a deep learning based method for automatic detection of MIZs in the RADARSAT-2 satellite images is presented. A synthetic aperture radar (SAR) dataset is manually annotated to train, test, and refine the method. Various convolutional neural network (CNN) models are evaluated as fixed feature extractors for the task of classification. To aid the classification accuracy a weighted binary cross-entropy loss criterion is used. Finally, to refine the segmentation process, a multi-scale patch technique is used. In this technique, for each image, different segmentation masks are calculated independently for different patch sizes. These segmentation masks are stacked together and averaged with equal weights. A thresholding operation is performed on this mask to generate the MIZ probability mask for the image. The analysis of the results demonstrates that CNN model predictions obtained with multiple sizes of spatial windows are able to detect MIZs in SAR images. The objectives of the study presented in this chapter are:

- To annotate and preprocess a SAR image dataset to be suitable as an input to the CNN.
- To test various CNN models as fixed feature extractors for MIZ classification.
- To build a segmentation pipeline using a trained classifier for MIZ detection.

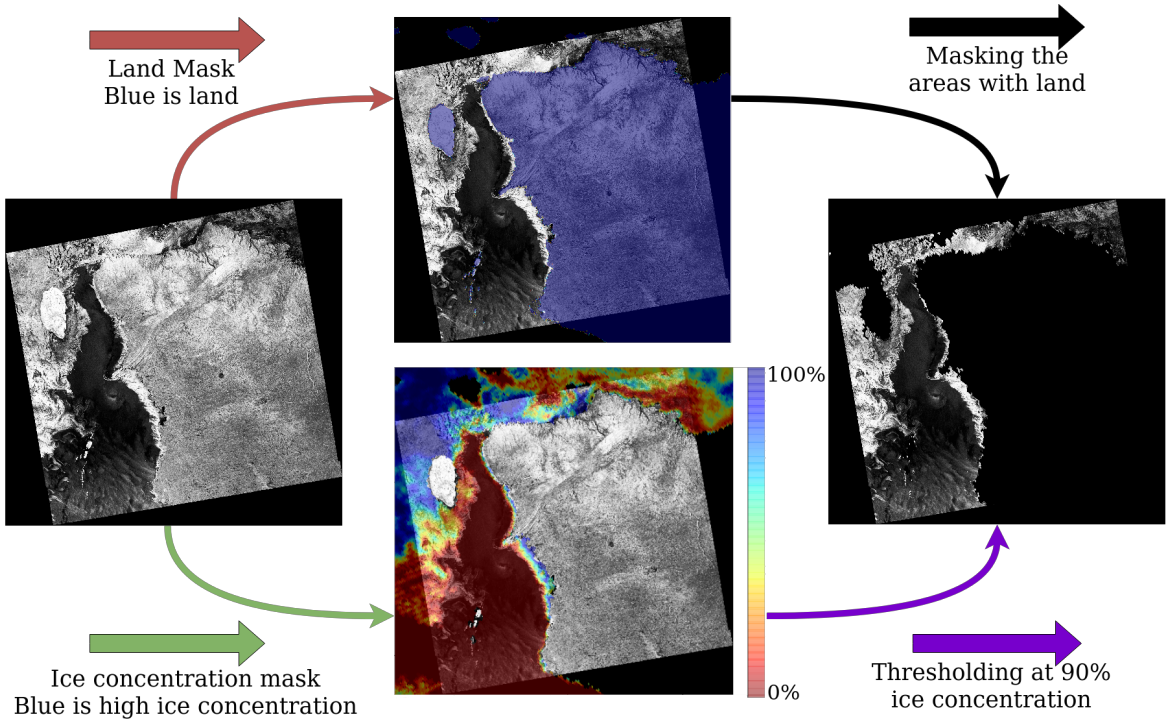


Figure 3.1: Data pre-processing pipeline for the segmentation step. Land and ice concentration masks are computed for every image. A thresholding operation is performed for both the masks. In our experiments, the threshold was set to 90% ice concentration for the ice mask.

3.1 Dataset

The dataset is comprised of C-band ScanSAR wide beam mode images acquired from RADARSAT-2 satellite in HH (horizontal transmit and receive). These images were down-sampled 4-fold from an original size of $10,000 \times 10,000$ pixels with an original spatial resolution of 50×50 m, by averaging over 4×4 blocks.

Images span over six ice seasons, covering the years 2008 to 2013. Each year the ice season begins in November and lasts until July of the following year. The geographical area under study is comprised of Hudson Strait located in Eastern Canada and its outflow into the Labrador Sea in the North Atlantic. The dataset is divided into two parts, classification, and segmentation.

The classification set was constructed by extracting patches from 160 images containing

MIZs spanning five seasons. For patch extraction, manual annotations were created which were verified by a domain expert. In total, there are 57 positive patches having MIZs and 493 negative patches (either open water (OW) or consolidated ice (CI)). The patch size varies from 36×36 pixels to 151×151 pixels.

The segmentation set contains 105 images from the 2008 season. Images from the segmentation set were masked using land and high ice concentration masks as shown in Fig. 3.1. Masking reduces the redundant data in the SAR images as MIZs are a water body phenomenon. The effective area of search is reduced to water bodies with less than 90% ice concentration.

3.2 Methodology

In this chapter, a two-step approach is employed. In the first step, a CNN network is trained for the classification task. The trained network is able to classify input patches as positive (MIZ) or negative (OW or CI). Densenet is selected as the classifier as it has a relatively low number of trainable parameters and higher image classification accuracy with respect to other classification models [44]. The best model is selected based on the validation accuracy and used later in the segmentation step. Images from the segmentation set are sliced into patches and fed to the classifier for classification. Later, the per patch classification scores are joined and scaled to the size of the images to render the MIZ probability masks.

3.2.1 Experimental Setup

For the experiments, PyTorch 1.3.1 and Torchvision 0.4.2 were used [80]. DenseNet161 architecture pre-trained on the ImageNet dataset [24] was selected because it had the best overall performance on our classification set as shown in Table 3.1. The final fully connected layer of the model was replaced with a new layer having 2 output nodes, a softmax activation was used to obtain the probability scores. All the layers except this new fully connected layer were frozen during training i.e., these parameters were not updated during the optimization stage. Thus, the network was used as a fixed feature extractor.

Table 3.1: Comparison of various deep learning models for patch classification step.

Method	Accuracy	Precision	Recall	F1-Score
AlexNet	0.86	0.39	0.25	0.30
VGG16	0.89	0.80	0.44	0.57
ResNet50	0.90	0.64	0.78	0.70
DenseNet161	**0.98**	**0.90**	**1.00**	**0.95**

3.2.2 Classification

The classification set was split into training and validation sets with an 80/20 ratio. The images were resized to $224 \times 224 \times 3$ since the DenseNet161 architecture is designed for RGB images of this input size. Adam optimizer was used with default settings and a learning rate of 10^{-4}. A batch size of 16 was used and weighted binary cross-entropy (BCE) was used as the loss function, which is defined below:

$$BCE = -\frac{1}{N}\sum_{i=1}^{N}\left[a\ y_i \log(\hat{y}_i) + b\ (1-y_i)\log(1-\hat{y}_i)\right] \tag{3.1}$$

where y_i is the label or ground truth and $\hat{y}_i$ is the prediction for the i^{th} sample, a and b are the weighting coefficients. For our experiments, a and b were chosen as 0.9 and 0.1 respectively. These numbers represent the ratio of the two classification classes. The best model was selected based on the highest validation accuracy and was used for the segmentation step. Data augmentation techniques namely rotation, flipping, and translation were used to increase the dataset size for training. In deep learning, data augmentation is a widely used technique for handling over-fitting which also helps to increase the generalization capability of the model.

3.2.3 Segmentation

The segmentation pipeline is shown in Fig. 3.2. Segmentation masks, representing the probability of MIZs were generated for the images in the segmentation set. Each segmentation mask is calculated independently for each image for different patch sizes, namely 60×60 pixels, 90×90 pixels, 120×120 pixels, and 180×180 pixels with a stride of half the patch size, respectively. Segmentation masks of each image are stacked together and averaged with equal weights, followed by thresholding at 50% confidence. The final

segmentation mask or MIZ probability mask is overlayed on the actual unmasked image to visualize the results.

3.3 Results and Discussion

Initially, classical machine learning algorithms such as logistic regression (LR) and one-class support vector machine (OCSVM) were investigated for patch classification, however unsatisfactory results were obtained. Hence, deep convolutional networks were explored for this task as they have demonstrated high classification accuracy on big image classification datasets. One of the reasons for their success is the ability to learn rich feature sets from the given samples. Deep convolutional networks such as AlexNet, VGG16, ResNet50, and DenseNet161 were compared and DenseNet161 outperformed all other methods for our classification set as shown in Table 3.1. Hence, DenseNet161 was chosen for the segmentation step. MIZ predictions over various SAR scenes are shown in Fig. 3.3, Fig. 3.3, Fig. 3.5, Fig. 3.6 and Fig. 3.7.

It was observed that the patch size of 90×90 pixels is highly successful in detecting the MIZs, in that it returned large probabilities over regions that appeared to be MIZs. This could be because both the mean and median of dimensions of the positive samples (MIZs) in the classification set used to train the model are 90 pixels (meaning most of the regions manually selected as corresponding the MIZs had a patch size of 90 pixels). An overall similar patch size might help to classify MIZs in the segmentation step. However, classifications on this patch size also generated some false positives, in particular over open water wind roughened regions as shown in Fig. 3.7. To overcome this problem, we intend to experiment with wind masks so as to mask the areas with high wind intensities. In the context of our multi-scale approach, we observed that the patch size of 180×180 pixels is effectively able to detect open water regions because it is easier for the classifier to make a decision when looking at a bigger area, but the segmentation masks obtained at this patch size lose localization. So, we chose a multi-patch size approach to decrease false positives while detecting most of the MIZs in the images which is evident from the qualitative analysis of the results. It is also observed that false-positive occurrence and inference time per image was reduced upon the application of land and ice concentration masks due to the reduction in the overall data to be processed per image.

In this chapter, an automated method to detect MIZs in RADARSAT-2 satellite images is presented. It is shown that a pre-trained CNN can easily classify the above phenomenon when used as a fixed feature extractor. When combined with a multi-scale patch approach, the fixed feature extractor is able to detect MIZs in the SAR images by taking advantage

of different sizes of spatial windows even in the case of a data shortage. The obtained classification results suggest that DenseNet161 performs better than other comparable CNN architectures for the task of MIZ classification in SAR images.

In the next chapter, we dive deeper into the marginal ice zones and focus on segmenting sea ice floes in SAR images.

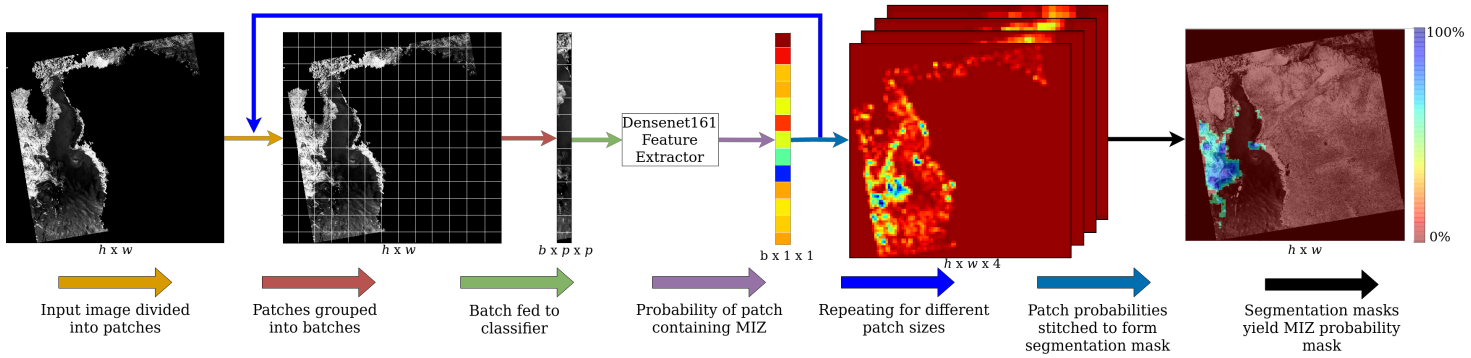
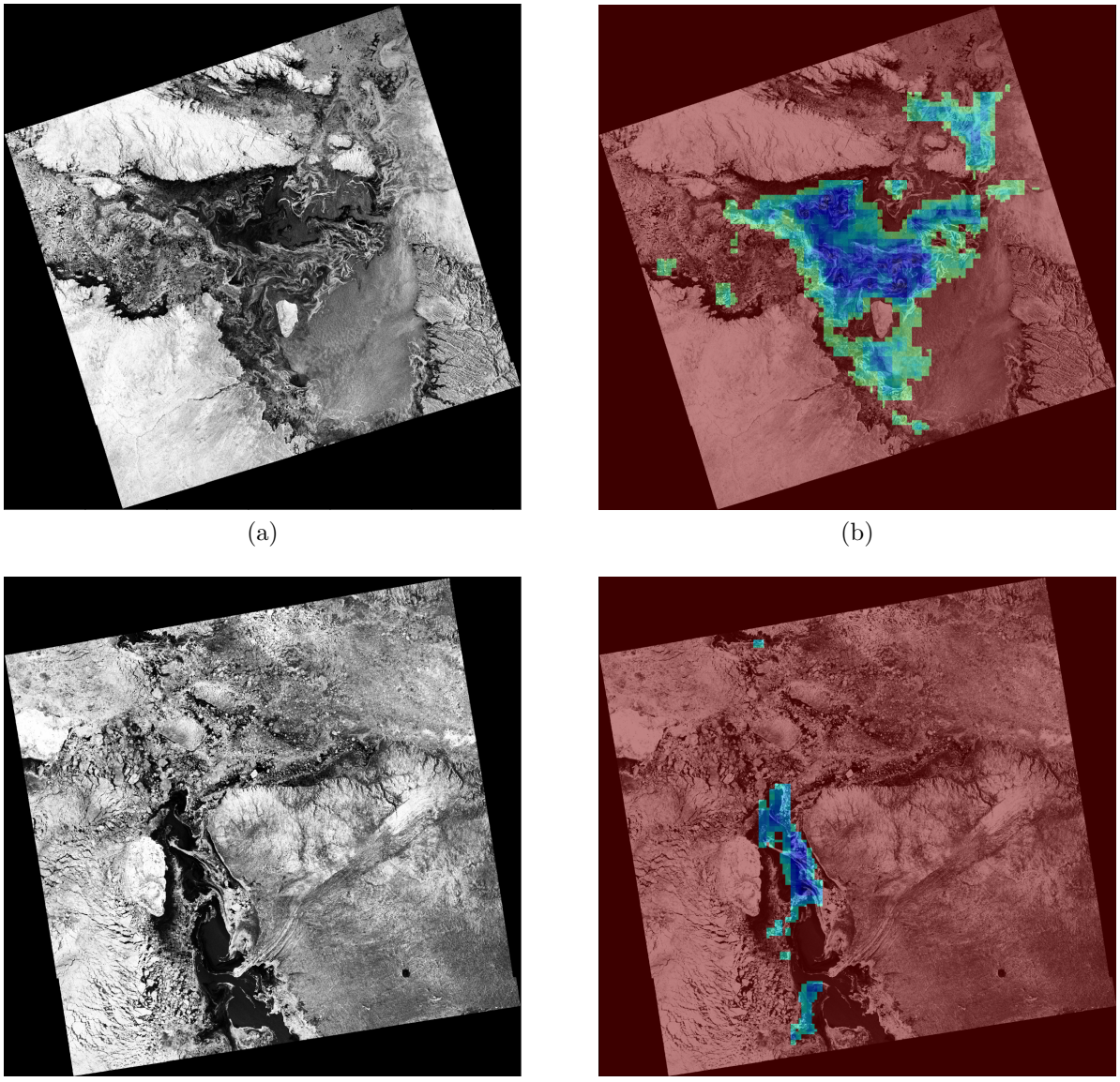


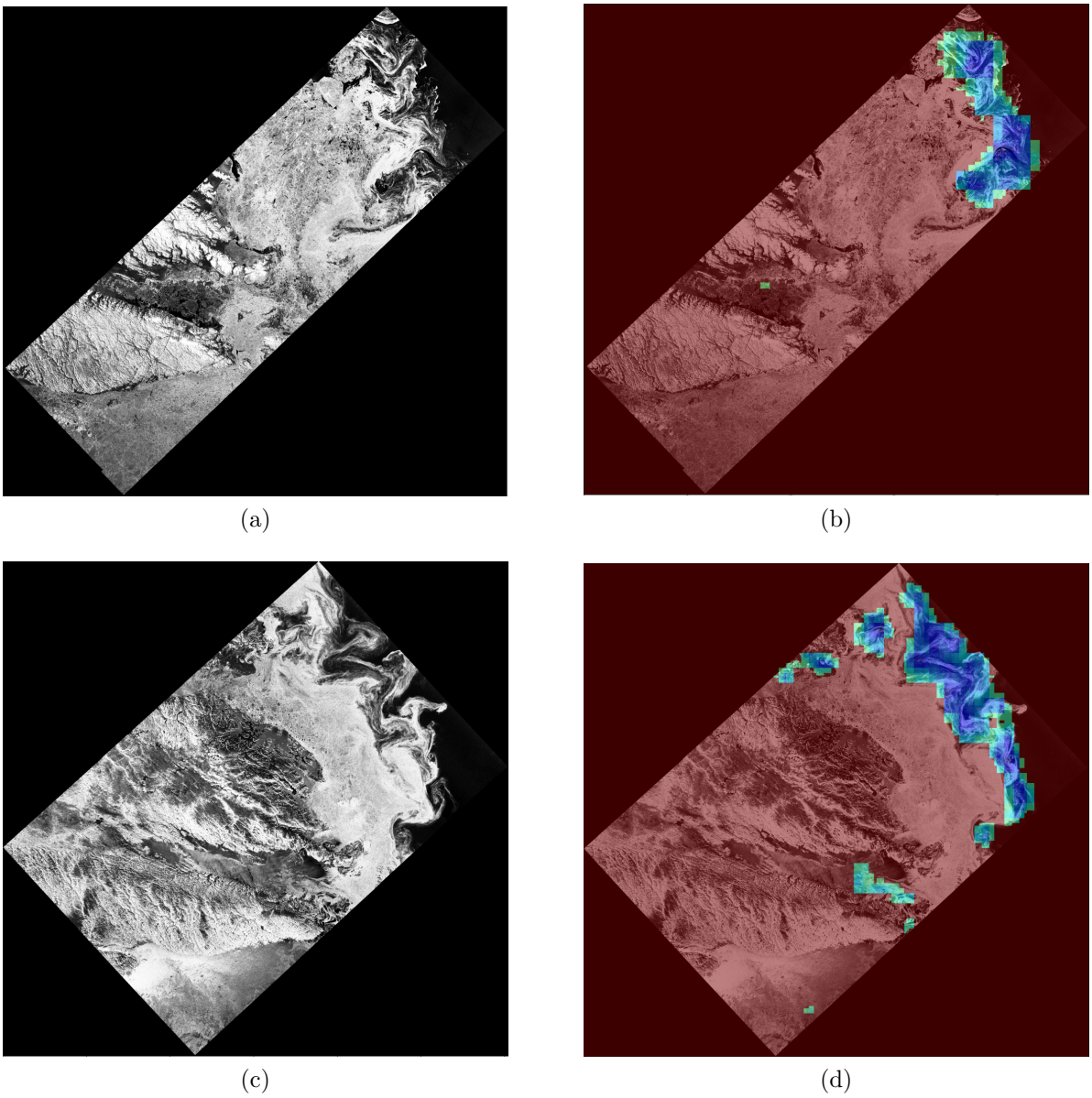
Figure 3.2: Figure shows the segmentation pipeline. First, the input image of size $h \times w$ is zero-padded on the sides. Next, the image is divided into a grid with $p \times p$ patch size. Batches created from slicing this grid are then fed to the neural network. For every patch, the output of the neuron corresponding to the MIZ class is assigned to the entire patch area to build a segmentation mask of size $h \times w$. This process is repeated for different patch sizes and the resultant masks are stacked together to get a tensor of size $h \times w \times 4$. Finally, a weighted average of this tensor is taken across the 3^{rd} dimension and followed by a thresholding operation to generate the MIZ probability mask.



(a) (b)

(c) (d)

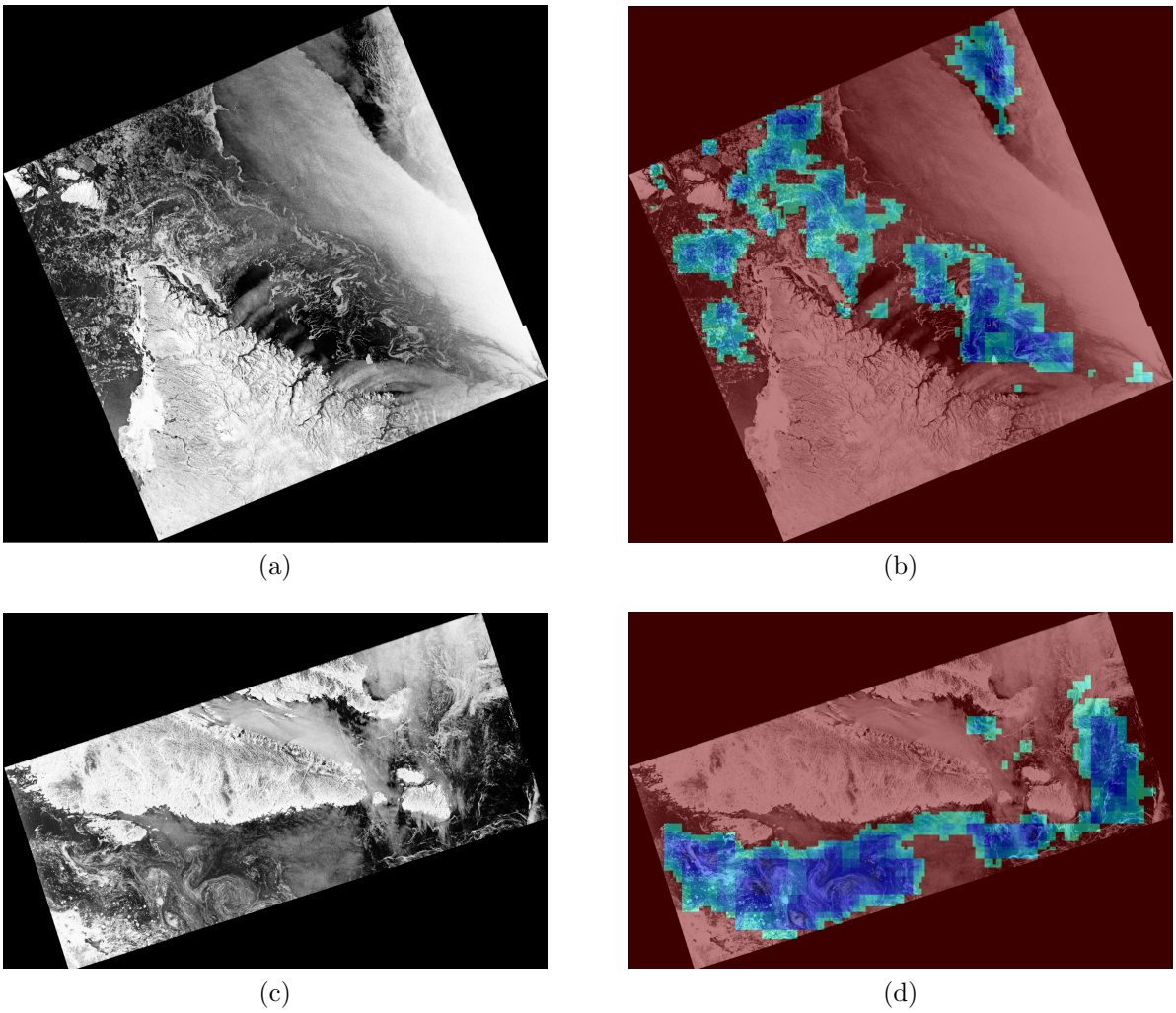
Figure 3.3: Sample segmentation outputs of some SAR scenes. (a) and (c) are SAR scenes with acquisition times 20091210105623 and 20091219113354 respectively. (b) and (d) are the respective MIZ prediction maps overlayed over the SAR images. High MIZ probability (Blue). Low MIZ probability (Red).



(a) (b)

(c) (d)

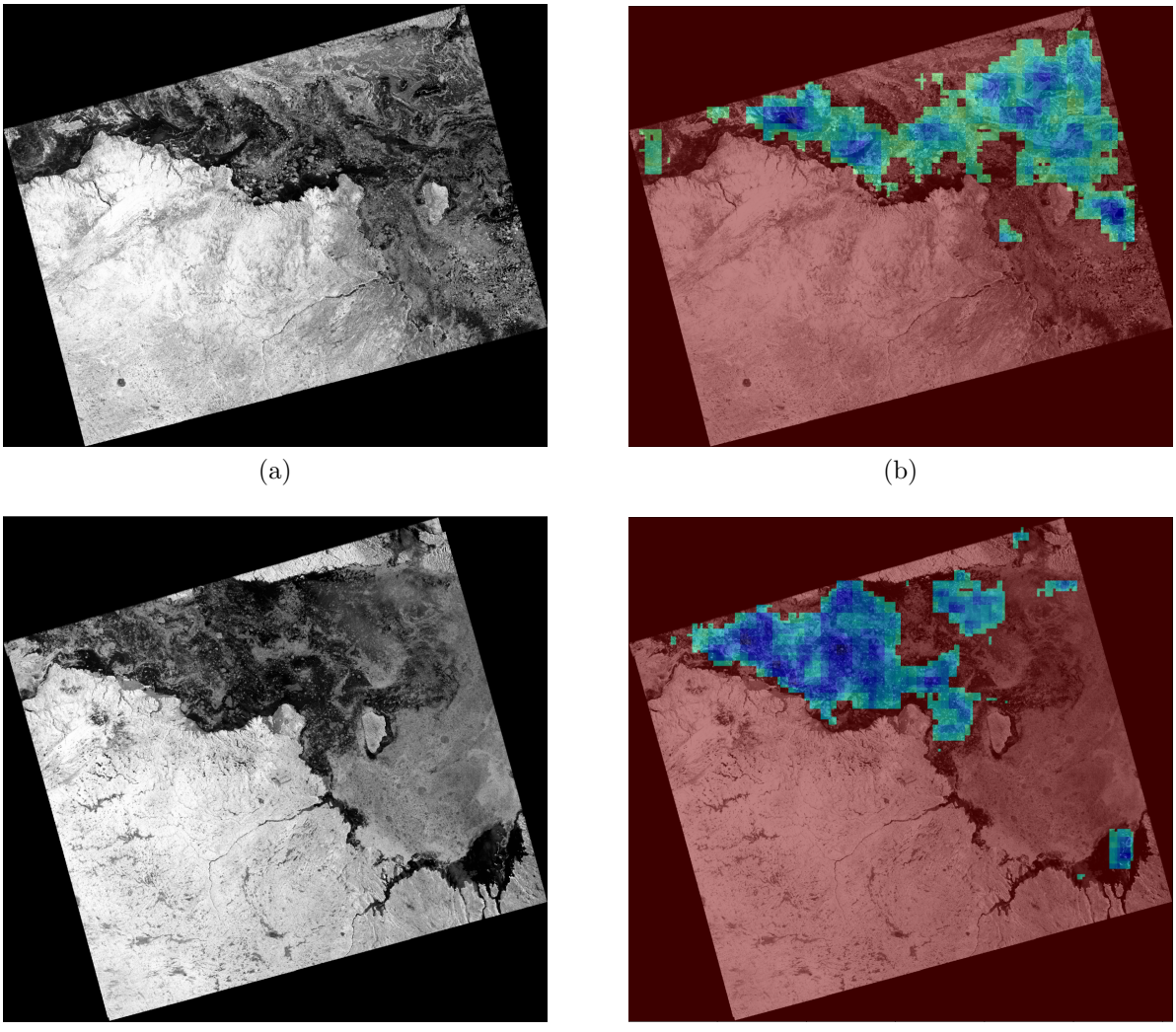
Figure 3.4: Sample segmentation outputs of some SAR scenes. (a) and (c) are SAR scenes with acquisition times 20100417215245 and 20100528215657 respectively. (b) and (d) are the respective MIZ prediction maps overlayed over the SAR images. High MIZ probability (Blue). Low MIZ probability (Red).



(a) (b)

(c) (d)

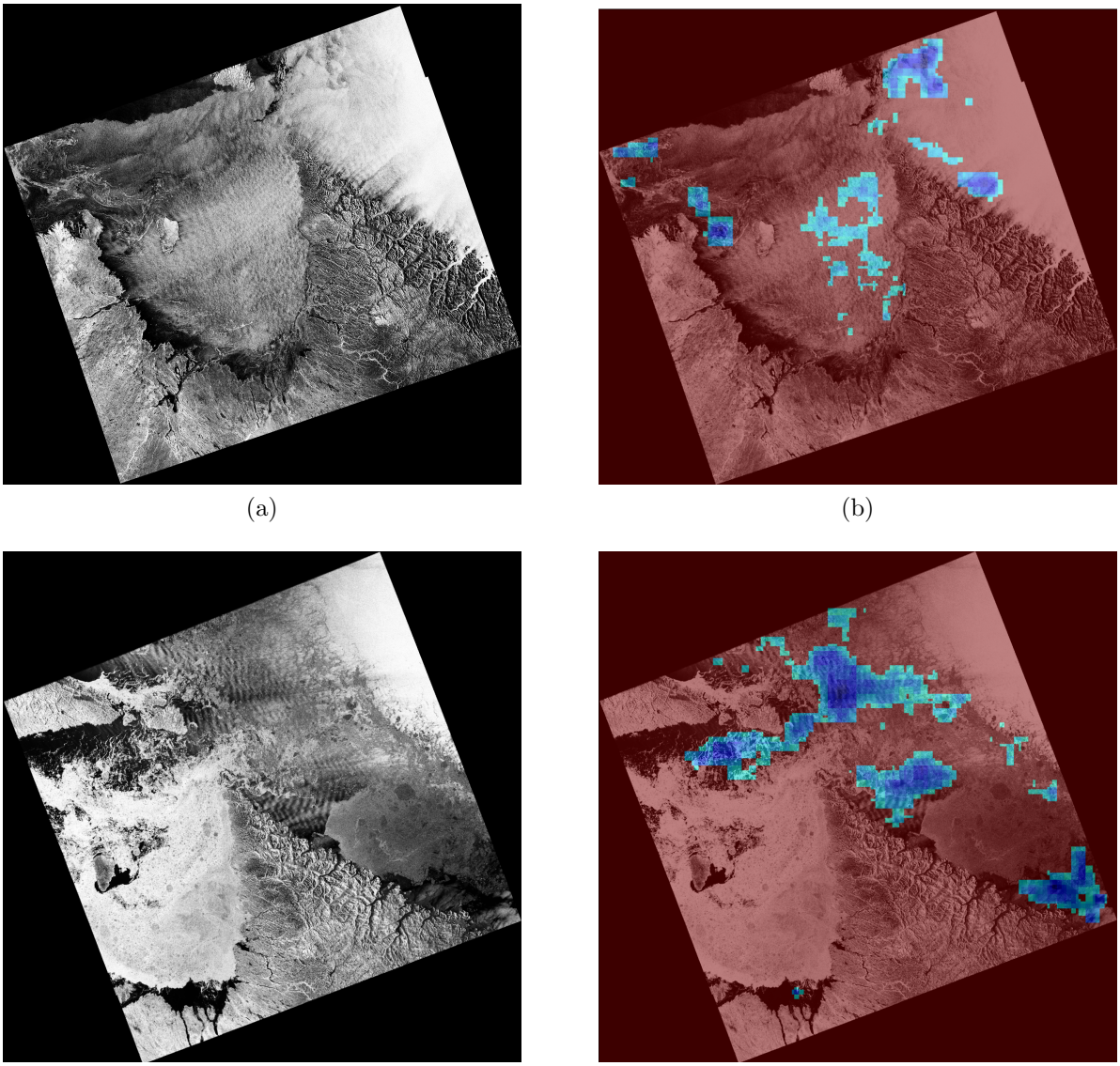
Figure 3.5: Sample segmentation outputs of some SAR scenes. (a) and (c) are SAR scenes with acquisition times 20110516103151 and 20110615105638 respectively. (b) and (d) are the respective MIZ prediction maps overlayed over the SAR images. High MIZ probability (Blue). Low MIZ probability (Red).



(a) (b)

(c) (d)

Figure 3.6: Sample segmentation outputs of some SAR scenes. (a) and (b) are SAR scenes with acquisition times 20121221110932 and 20140609110514 respectively. (b) and (d) are the respective MIZ prediction maps overlayed over the SAR images. High MIZ probability (Blue). Low MIZ probability (Red).



(a) (b)

(c) (d)

Figure 3.7: Examples of SAR scenes with relatively low classification accuracy where wind roughened waters are classified as MIZs. (a) and c (scene acquisition times 20081205104812 and 20140603104006 respectively) are SAR scenes with wind roughened ocean water. (b) and (d) are the respective MIZ prediction maps overlayed over the SAR images. High MIZ probability (Blue). Low MIZ probability (Red).

Chapter 4

RUF: A model for sea ice floe segmentation

In the last chapter, a marginal ice zone (MIZ) segmentation pipeline using transfer learning was explained. The proposed method used a patch wise approach where classifications over a small patch size were combined to generate the segmentation mask for the whole image. In the method, the same is processed 4 times to obtain the desired prediction mask. Although the generated masks show good results upon verification by visual inspection, the time spent in processing one image is quite large due to the reprocessing involved.

In this chapter, we propose a segmentation model tailored for detecting ice floes in SAR images. Rather than using a classifier for patch classification and later stitching the individual outputs, we use a fully convolutional network that generated the output segmentation masks of the same size of the input image. The model exploits the advantages of both convolutional neural networks and convolutional conditional random field (Conv-CRF) in a combined manner. The residual UNET (RES-UNET) computes expressive features to generate coarse segmentation maps while the Conv-CRF exploits the spatial co-occurrence pairwise potentials along with the RES-UNET unary/segmentation maps to generate final predictions. The whole pipeline is trained end-to-end using a dual loss. Comparison of experimental results with the conventional segmentation networks such as UNET, DeepLabV3, and FCN-8 demonstrates the effectiveness of the proposed architecture.

Figure 4.1: The geographical area of interest of captured SAR scenes. Red polygon represents the geographic extent of a SAR image used in the study.

4.1 Dataset

The dataset is composed of 9 RADARSAT-2 C-band ScanSAR wide-beam mode images acquired in HH (horizontal transmit and horizontal receive) polarisation. The images were acquired at a center frequency of 5.405GHz with a 500Km swath width and provide a nominal pixel spacing of 50m. The SAR images were captured from the area as shown in Fig. 4.1, with the red polygon describing the extent of one SAR image. Information regarding the image acquisition dates, central latitude, and longitude, instances of annotated floes is given in Table 4.1.

Table 4.1: Description of RADARSAT-2 scenes dataset that was used to train, validate, and test various model architectures.

Dataset	**Scene Acquisition**		**Number of Floes**	**Central coordinate**
	Date	Time		
Train	2008-11-13	11:29:00	204	63°32′40″N, 75°22′25″W
	2011-06-12	22:12:53	208	60°28′50″N, 68°47′22″W
	2012-06-05	22:42:59	177	62°38′24″N, 77°2′1″W
	2013-06-03	11:25:54	348	62°47′46″N, 72°38′1″W
	2013-06-14	22:34:35	203	62°38′14″N, 74°56′25″W
Total			**1140**	
Validation	2012-05-08	11:29:49	212	63°32′40″N, 75°22′25″W
	2012-06-06	10:44:18	128	60°42′8″N, 65°1′W
Total			**340**	
Test	2011-06-15	10:56:38	59	62°20′26″N, 67°29′54″W
	2014-06-15	11:29:40	288	63°30′41″N, 75°22′W
Total			**347**	

4.1.1 Data Preprocessing

SAR images have a grainy 'salt and pepper' appearance also called speckle noise that is caused due to random interference between coherent returns. To reduce this intrinsic contamination of speckle noise, the SAR images were down-sampled fourfold. The down-sampling operation was carried out by averaging over 4×4 pixel non-overlapping blocks.

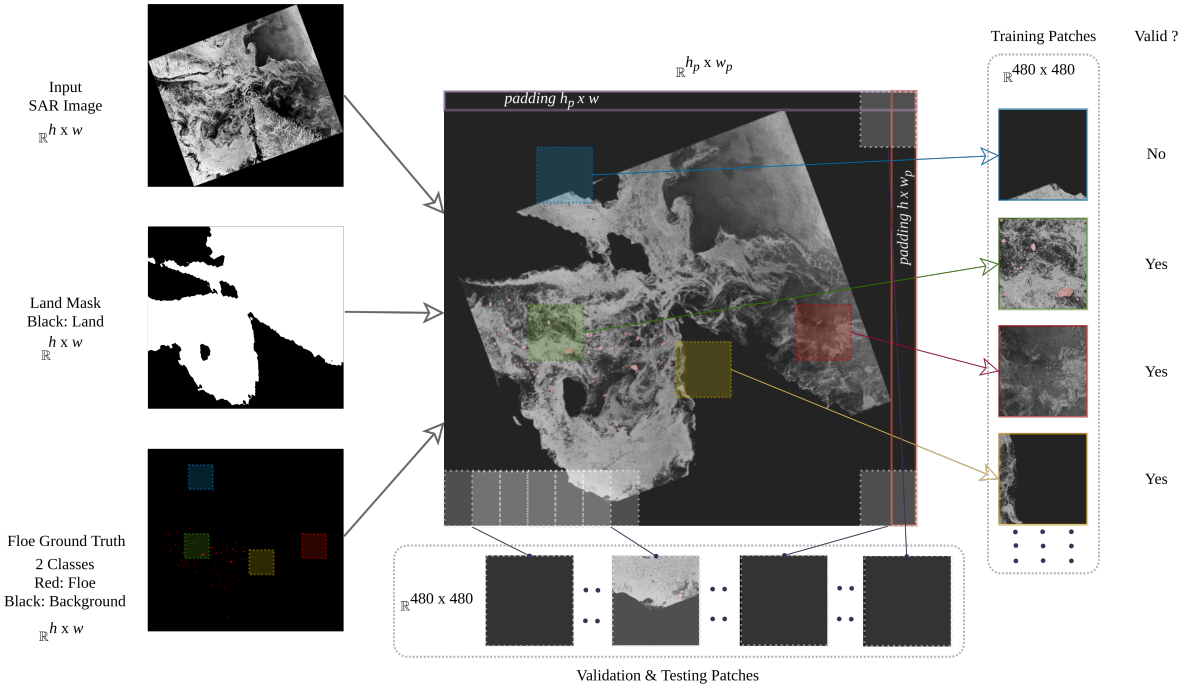


Figure 4.2: Patch generation for training, validation and testing. Training: input SAR image ($h \times w$) is land masked and patches of size 480×480 are selected randomly from the image. For a patch to be considered as a valid training sample, it should not contain more than 50% black pixels. Validation and testing: land masked SAR image is padded using black pixels. Patches are drawn serially from the land masked SAR image with a patch overlap of 50%. Total training/validation extracted from an image are $(h_p \times w_p) \div (480 \times 480)$

Downsampling operation changed the nominal pixel spacing to $200m$ with a reduction in data volume to 1/16 of the original.

Note that the Hudson Strait and the neighboring geographical areas of interest have a long coastline, and the images have a significant land cover. As ice floes are a water body phenomenon, pixels representing land in the images were masked as black as shown in Fig. 4.2. To generate the land masks for the images, we applied a threshold of $0m$ elevation to the elevation masks and later a gaussian blur with $(5, 5)$ kernel size to remove the rough edges along the sea-land boundaries. The elevation masks were generated using $ACE2\_5Min$ digital elevation model with bilinear interpolation using the European Space Agency (ESA) toolbox, Sentinel Application Platform.

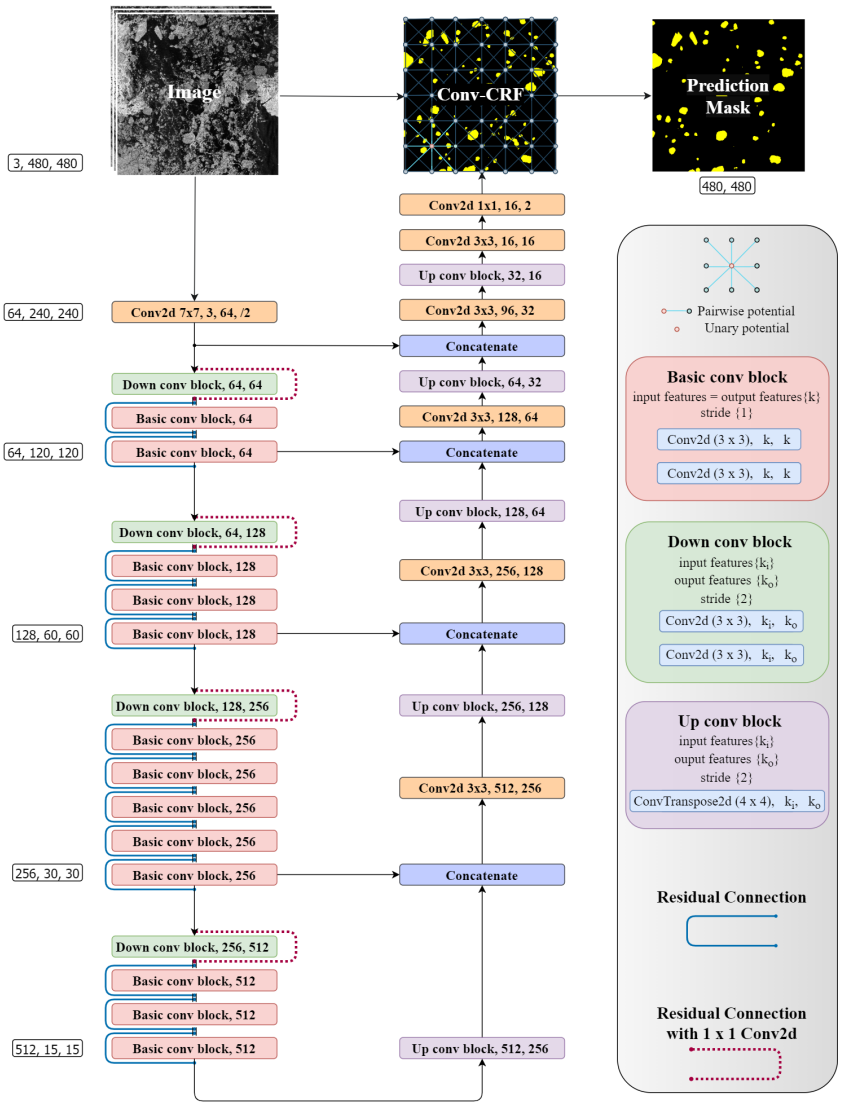


Figure 4.3: Figure of the proposed network architecture. The lower branch of the architecture consists of the downsampling/encoding path. The upper branch of the architecture consists of the upsampling/decoding path. Skip connections (interconnections between the two paths) help in integrating the location information in the downsampling paths to the contextual information in the upsampling path. The output of the final layer (loosely labeled mask) in the upsampling path is fed to the Conv-CRF as unary. Conv-CRF processes the loosely labeled mask along with input image to generate the Prediction Mask.

4.1.2 Data Annotation

The dataset contains 9 images with $1,627$ manual annotations of ice floes. For our experiments a single class was defined for annotation, namely 'floe' and all the remaining pixels were categorized as background. In the pixel-level annotations, we used the following criteria to annotate a closed contour as a floe:

- The contour contains consolidated ice as determined via visual inspection.
- At-least 30 percent of the contour boundary is in contact with seawater.
- The contour contains at-least 60 pixels.

These criteria were chosen to eliminate the closed contours in ice covers (frozen in floes) and to reduce noise artifacts. The floes were annotated through visual inspection of the imagery using the Computer Vision Annotation Tool (CVAT) [94].

4.2 Methodology

This section introduces the proposed RUF network architecture. The proposed model leverages the advantages of residual CNNs and Convolutional Conditional Random Fields (Conv-CRFs) [106] alongside dual loss function. Residual blocks allow a network to train easily while the UNET skip connections help in easy information propagation between different layers of the network. To facilitate the learning of the network weights, we jointly train the RES-UNET and Conv-CRF parts with the dual loss function. The overall architecture of the proposed network is illustrated in Fig. 4.3. Information about various components of the network is given in the subsequent subsections.

4.2.1 End-To-End training

For training the proposed architecture, we first feed the SAR images to the base RES-UNET network, where pixel-wise segmentation maps from the base network are fed to the secondary Conv-CRF network as the unary. The secondary Conv-CRF network cleans up spurious predictions and enhances object boundary predictions. Training these two parts in an end-to-end manner allows the gradients to flow through the whole pipeline and enables both the network to learn simultaneously. Hence, with this approach, we optimize both the models with respect to each other to provide optimum results.

4.2.2 Dual loss function

For the task of multi-class classification, network weights are generally trained using the categorical cross entropy loss. In case of segmentation, losses involving ground truth and prediction overlap are generally employed. To train our network, we optimize the proposed RUF architecture using a weighted dual loss function including Binary Cross Entropy (BCE) loss and Soft Dice (SD) loss:

$$\mathcal{L}_{total} = \alpha \cdot \mathcal{L}_{BCE} + (1 - \alpha) \cdot \mathcal{L}_{Dice}, \tag{4.1}$$

where $\alpha \in [0, 0.5, 1]$ is the weight parameter. The combination of BCE and Dice loss has been previously used for medical image segmentation [62]. For the task of ice floe segmentation, we have adapted the same with alpha as a weighted parameter. BCE loss measures the classification accuracy of the model prediction and it increases as the prediction diverges from the ground truth [119]. SD loss, which is derived from Dice Coefficient, measures the similarity between two sets [105]. BCE loss and SD loss can be defined as Eq: 4.2 and Eq: 4.3 respectively:

$$\mathcal{L}_{BCE} = -\sum_{i=1}^{N} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)], \tag{4.2}$$

$$\mathcal{L}_{Dice} = 1 - \frac{2 \sum_{i=1}^{N} y_i \hat{y}_i}{\sum_{i=1}^{N} \hat{y}_i + \sum_{i=1}^{N} y_i}, \tag{4.3}$$

where y_i is the label or ground truth and, $\hat{y}_i$ is the prediction for the ith pixel.

4.3 Metrics

The following metrics were used to evaluate the proposed approach and perform a comparison with other standard segmentation approaches.

1. **Mean Intersection over Union (mIoU)**: mIOU is a popular metric that is often used for the task of semantic segmentation. mIoU is calculated by averaging the Jaccard Score (J) over all the given classes in a segmentation task. Jaccard Score represents the ratio of area of intersection between the ground truth (G) and the

predicted segmentation (P) maps to the union between ground truth and predicted segmentation maps:

$$J(G, P) = \frac{|G \cup P|}{|G \cap P|}, \tag{4.4}$$

where G and P are the ground truth and predicted segmentation maps, respectively.

2. **Mean Pixel Accuracy (mPA):** mPA is the ratio of correctly classified pixels per class, then averaged over the total number of classes. For $k+1$ classes (k foreground classes and one background):

$$mPA = \frac{1}{k+1} \sum_{i=0}^{k} \frac{p_{ii}}{\sum_{j=0}^{k} p_{ij}}. \tag{4.5}$$

3. **F1 Score**: F1 Score is the harmonic mean of Precision and Recall. Precision and Recall:

$$Precision = \frac{TP}{TP + FP}, Recall = \frac{TP}{TP + FN}, \tag{4.6}$$

where TP is the number true positives, FP is the number of false positives and FN is the number of false negatives.

4. **Dice Score**: Dice Coefficient is the ratio of twice the area of intersection between ground truth and predicted segmentation maps, with the total number of pixels in the maps:

$$Dice = \frac{2|G \cap P|}{|G| + |P|}. \tag{4.7}$$

Both Dice Coefficient and IoU are positively correlated. In the case of binary segmentation where foreground class is considered as the positive class, F1 Score and Dice Coefficient are identical:

$$F1 = \frac{2|G \cap P|}{|G| + |P|} = Dice. \tag{4.8}$$

Hence, for our binary segmentation problem of ice floes, in this approach we use refer Dice Coefficient as F1 Score thereafter.

4.4 Experiments and Results

RUF is a five-level deep convolutional network with symmetric downsampling and upsampling paths as shown in Fig. 4.3. The downsampling path encodes the image into condensed representation while the upsampling path decodes this information into pixel-wise categorization. The downsampling path has four residual blocks. Each residual block contains multiple residual units built with two 3×3 convolutional layers and a residual connection. The convolutional layers are accompanied by a BatchNorm2d layer with a ReLu activation function. Rather than employing a max-pooling operation to downsample the feature maps [87], we use down conv blocks with strided convolutions. The upsampling path has a similar structure but rather than using residual blocks we use simple convolutional layers. Feature maps are upsampled at each level using up conv blocks employing transposed convolutions. Skip connections concatenate the output of each level in the downsampling path to the upsampling path and helps in combining coarse information with finer information. At the final level, a 1×1 convolution projects the multi-channel feature maps to our intermediate segmentation mask. This mask is then processed in conjunction with the input image to calculate the unary and pairwise potentials of the Conv-CRF for further refinement. A softmax operation is applied to the output of Conv-CRF to obtain, which is later thresholded at 50 percent confidence to obtain the prediction mask.

4.4.1 Dataset Split

We split the dataset into train, validation, and test sets with a rough $60 : 20 : 20$ ratio split, such that the train set contains 60% of annotated floes while validation and test sets each contain 20% of floe annotations. Full SAR images were placed in each set in order to have truly independent samples across different sets, as shown in Table 4.1. Splitting the dataset as such allows us to ascertain better generalization of the model.

Training Procedure

There are a total of 1140 annotated floes in the train set which comprise approximately 60% of the overall annotated floes. Due to limited training images, we employ a random patch draw policy to train the model. An image is first randomly selected from the train set. We then randomly draw a patch of the given patch size from this selected image. Fig. 4.2 illustrates the training patch selection process. A patch is considered a valid training sample if it does not contain more than 50% black area (due to image boundaries

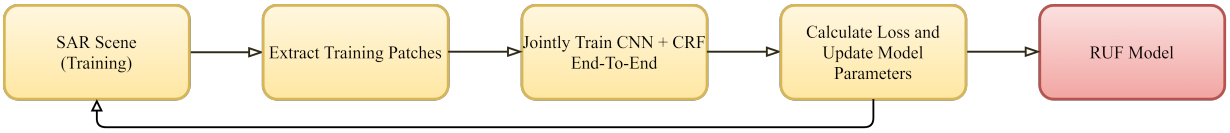


Figure 4.4: Model training pipeline. Input SAR image is processed as illustrated in Fig. 4.2 to extract patches for the training batch. These patches are fed to the model and the model parameters are updated according to the loss.

Figure 4.5: Model testing pipeline. Input SAR image is processed as illustrated in Fig. 4.2 to feed patches serially to the trained RUF model for inference. These patches are then reconstructed to yeild the segmentation mask for the whole image.

or land masking) or contains at least one floe either fully or partially. For example, a patch with 70% land and containing one floe is a valid sample. The above process is repeated until we find enough training samples for the training batch. We randomly rotate the eligible training samples by 0, 90, 180, or 270 degrees as a data augmentation technique to increase the overall data available to train the models. This data augmentation provides a regularisation effect that helps the models to generalize better on the overall dataset by reducing over-fitting.

Validation & Testing Procedure

The validation and test sets contain 2 images each, which account for approximately 20% of the overall annotated floes. To validate and test the model, patches are extracted serially from the images with an overlap of 50%. Fig. 4.2 provides details about selecting validation and testing patches. The validation dataset is used to check if the model is overfitting while the test dataset is used to compare different models.

4.4.2 Setup

The code was implemented using Pytorch 1.3.1 and Torchvision 0.4.2 open source frameworks. All the experiments were performed on a Ubuntu 18.04 LTS system with an Intel i7-4790 CPU@3.6 GHz coupled with 32GB DRAM while using a single NVIDIA GTX

Table 4.2: Comparison of different unet backbones on validation set. Best results are in **Bold**.

Model	Parameters	mIOU	F1	mPA
UNET	7,788,545	70.50	58.77	74.83
VGG-UNET	20,267,281	71.52	60.71	73.62
Inception-UNET	27,161,264	71.59	60.89	74.68
RES-UNET	25,658,402	**72.57**	**62.88**	**80.40**

TITAN X (PASCAL) with 11GB VRAM, running NVIDIA driver 440.64.00 and CUDA 10.2 for computational tasks.

For all experiments, model weights were initialized using kaiming_uniform initialization [36]. To optimize the model parameters further, we used the ADAM optimizer at an initial $1e-4$ learning rate. The Conv-CRF was initialized using default parameters from [59, 106] with an exception, We removed the Gaussian blur as the SAR images in our dataset are already downsampled with 4-fold averaging.

4.4.3 UNET Backbone-Selection

To select the primary segmentation network for our pipeline, we first compared the UNET architecture with different backboned UNETs. The key difference between a UNET and a backboned-UNET being that the two convolutional layers and a 2×2 maxpooling operation for each level in the downsampling path are replaced with the different convolutional blocks of the backbone architecture. While the skip connections and the convolutional layers in the upsampling path remain the same. Information regarding the number of parameters and comparison between different UNET architectures is given in Table 4.2. We use VGG19, Inception V3, and ResNet34 to construct different UNET architectures. We observe that UNET architecture with a ResNet34 based encoder achieves the best scores on the validation set and is selected for further improvement. Refer to Table 4.2 for more information.

4.4.4 Joint training with Conv-CRF

With the UNET backbone selected, we examine different configurations of joint training. We experiment with decoupled learning, stepwise learning, and end-to-end learning ap-

Table 4.3: Comparison of different loss functions for various model architectures using end-to-end training approach on validation set. BCE: Binary cross entropy loss; Dice: Dice loss; BCE + Dice: Equally weighted BCE and Dice Loss. Best results are in **Bold**.

Model	**Loss function**								
	BCE			**Dice**			**Dice + BCE**		
	mIoU	F1	mPA	mIoU	F1	mPA	mIoU	F1	mPA
UNET	65.76	48.66	66.39	65.64	48.61	71.96	70.51	58.78	74.83
RES-UNET	70.83	59.40	74.41	70.32	58.44	75.41	72.57	62.88	80.40
RUF (*Ours*)	70.91	59.59	75.36	71.39	60.48	74.61	**75.07**	**67.28**	**81.83**

Table 4.4: Comparison of different RUF joint training approaches on validation set. Best results are in **Bold**.

Method	**mIoU**	**F1**	**mPA**
Decoupled	68.25	54.23	72.90
Stepwise	69.73	57.27	74.87
End-to-end	**75.07**	**67.28**	**81.83**

proaches. The results of which are illustrated in Table 4.4. Decoupled learning is based on the assumption that CRF needs an accurate prediction of the unaries [59] to learn efficiently. In this approach initially only the CNN model is training till loss convergence, i.e. when the loss begins to saturate. After this step, the CRF is trained as a standalone model with the CNN output as unary to the CRF. The gradients never flow through the whole model in an iteration. Stepwise learning is similar to decoupled learning but in the second step, both the CNN and CRF are trained jointly such that weights of the whole model are updated in the second step. End-to-end learning involves training the whole architecture jointly from the very beginning such that the CNN and CRF are able to co-adapt.

We observe that for our task the end-to-end learning approach yielded preeminent results and thus was used for further experiments.

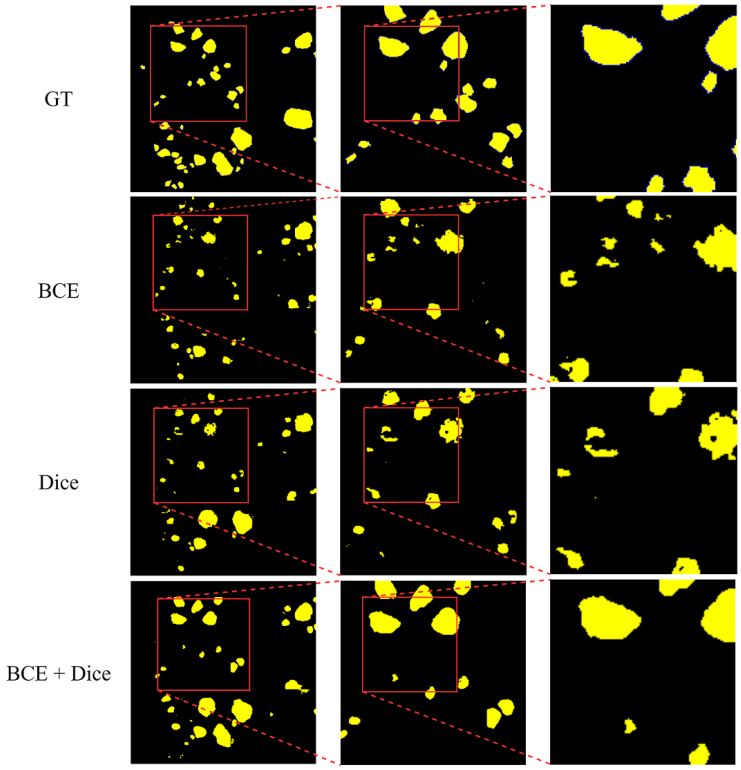


Figure 4.6: Comparison between the predictions of RUF model trained using BCE, Dice and BCE + Dice loss. It can be observed that the predictions obtained by using BCE + Dice loss have discrete continuous boundaries similar to the ground truth (GT).

4.4.5 Dual loss function selection

To select the optimum weights of our dual loss function, we trained different RUF models using various α values for $\mathcal{L}_{total}$ in Eq. 4.1. We evaluated three different configurations with $\alpha \in [0, 0.5, 1]$, which yield only BCE loss, equally-weighted BCE and Dice loss, and only Dice loss respectively. The results of our experiments are given in Table 4.3. We observe that using a dual loss comprising of equally-weighted BCE and Dice loss helps the network to generate better predictions. We observe that training with either only BCE or Dice loss does help the model loss to converge quickly but the predictions from such models lack discrete boundaries as illustrated in Fig. 4.6.

Table 4.5: Comparison of different models for various patch sizes on validation set. Best results are in **Bold**.

Patch Size	Batch Size	mIoU	F1	mPA
96	64	70.53	58.81	74.09
192	32	70.77	59.38	77.65
288	16	70.75	59.24	74.16
384	12	71.99	59.77	74.40
480	8	**75.07**	**67.28**	**81.83**
576	6	69.85	57.44	72.13

4.4.6 Patch size selection

As the overall segmentation network configuration is selected, we investigate further to choose the optimum patch size to train our model. We trained various RUF models beginning with a patch size of 96×96 pixels with a batch size of 64 patches. We observed that this model was able to identify small floes, while it struggled to detect floes which appear partially in the patch. To overcome this issue we decided to increase the patch size until we obtain optimum results. To train the subsequent models with bigger patch sizes, we decreased the batch size to obtain optimum GPU memory utilization. Results of our experiment are given in Table 4.5. We observed that there is a direct correlation between bigger patch size and the segmentation accuracy as evident from increasing metrics in Table 4.5, with an exception to the patch size of 576×576 pixels. One of the reasons for this reduction is a low batch size. Overall, we observed that the patch size of 480×480 resulted in the best performance.

4.4.7 Comparison with other models

Using the optimal input size and other network parameters, we obtained the best results using our proposed method. We then compared our method with several other standard segmentation models that are typically used in SAR image segmentation. Test dataset images are used in this comparison. To gauge the iterative improvement component of our method, we plot the predictions of UNET, RES-UNET, and RUF together in Fig. 4.7. Upon initial observation, we notice that both UNET and RES-UNET generate a greater number of false-positive predictions as compared to RUF. We observe that RUF is able

Table 4.6: Comparison between various segmentation architectures on test set. Best results are in **Bold**.

Model	**Parameters**	**mIoU**	**F1**	**mPA**
UNET	7,788,545	69.20	56.54	71.87
RES-UNET	27,912,770	71.76	61.24	74.66
FCN8	54,304,086	69.88	57.48	72.24
DLV3	60,991,062	67.47	52.43	68.94
RUF (ours)	26,119,208	**74.52**	**66.26**	**77.88**

to detect floes of various shapes and sizes, while UNET struggles to detect bigger floes. UNET is also unable to delineate the floes in close proximity to each other. The main reason for this issue would be the absence of residual connections in the UNET, as the model is unable to learn finer features as compared to other more complex models. It can also be observed that both UNET and RES-UNET generate multiple broken boundary predictions where a single floe is segmented as multiple floes.

We also compare RUF with other standard segmentation models such as FCN-8 (FCN) and DeepLabV3 (DLV3), the predictions of which are illustrated in Fig. 4.8. We can observe that the predictions from both FCN and DeepLab do not suffer from false positive, however, they do struggle to detect bigger floes and delineate floes in close proximity. The absence of a dense component such as a conditional random field to enhance finer segmentation details would explain this issue. Apart from these qualitative observations, the superiority of RUF for the task of ice floe segmentation is also evident from the quantitative analysis as given in Table 4.6.

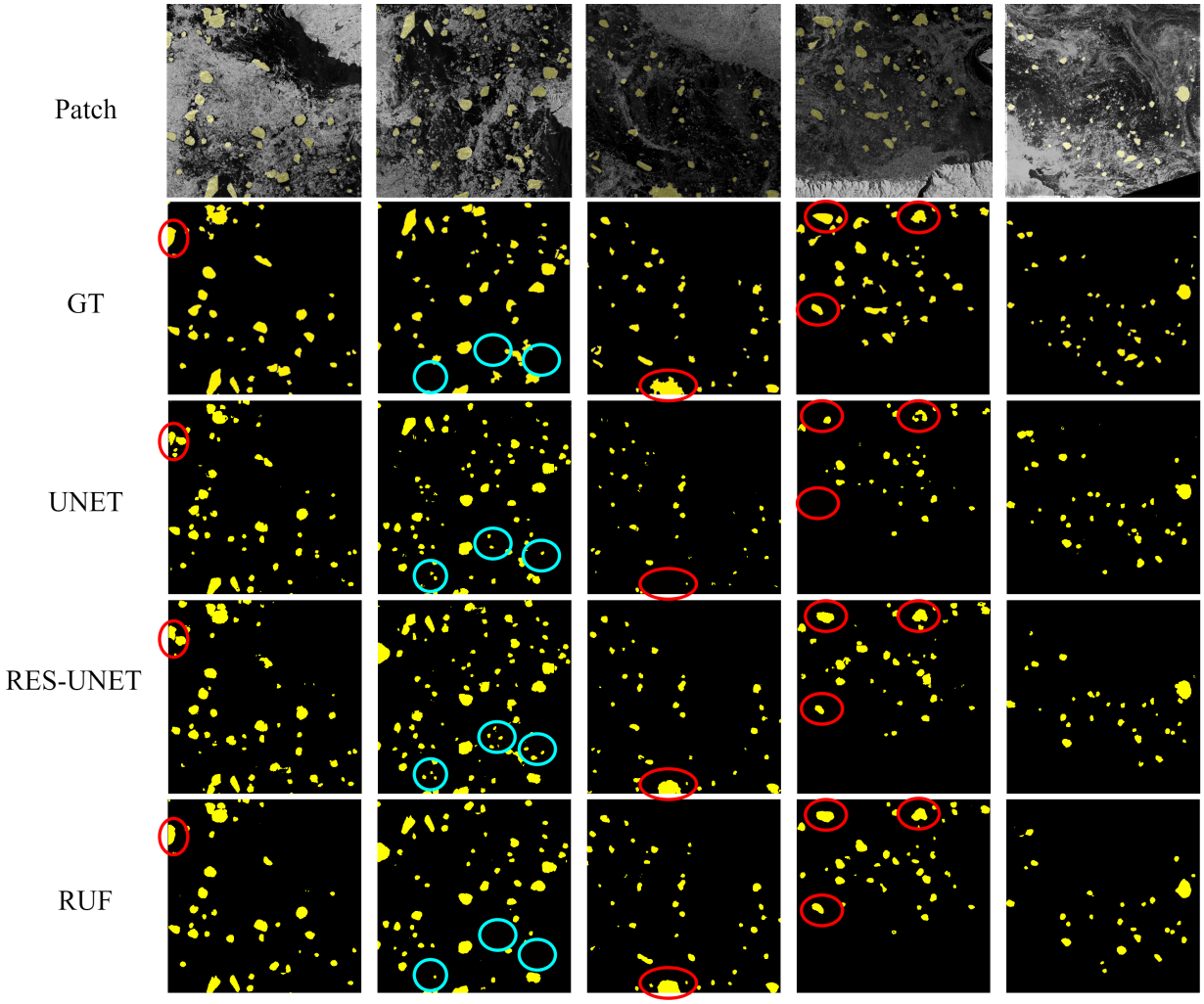


Figure 4.7: Comparison between the predictions of UNET, RES-UNET, and RUF for the task of ice floe segmentation on patches from the test set. Red ovals highlight some segmentation results where both UNET and RES-UNET are unable to properly segment a fully or partially visible ice floe. Blue ovals highlight some segmentation results where both UNET and RES-UNET give many false positive predictions. It can be observed that the proposed RUF architecture is able to produce finer segmentation results as compared to other backbone architectures. Here, Patch and GT denotes original image patch input and ground truth respectively.

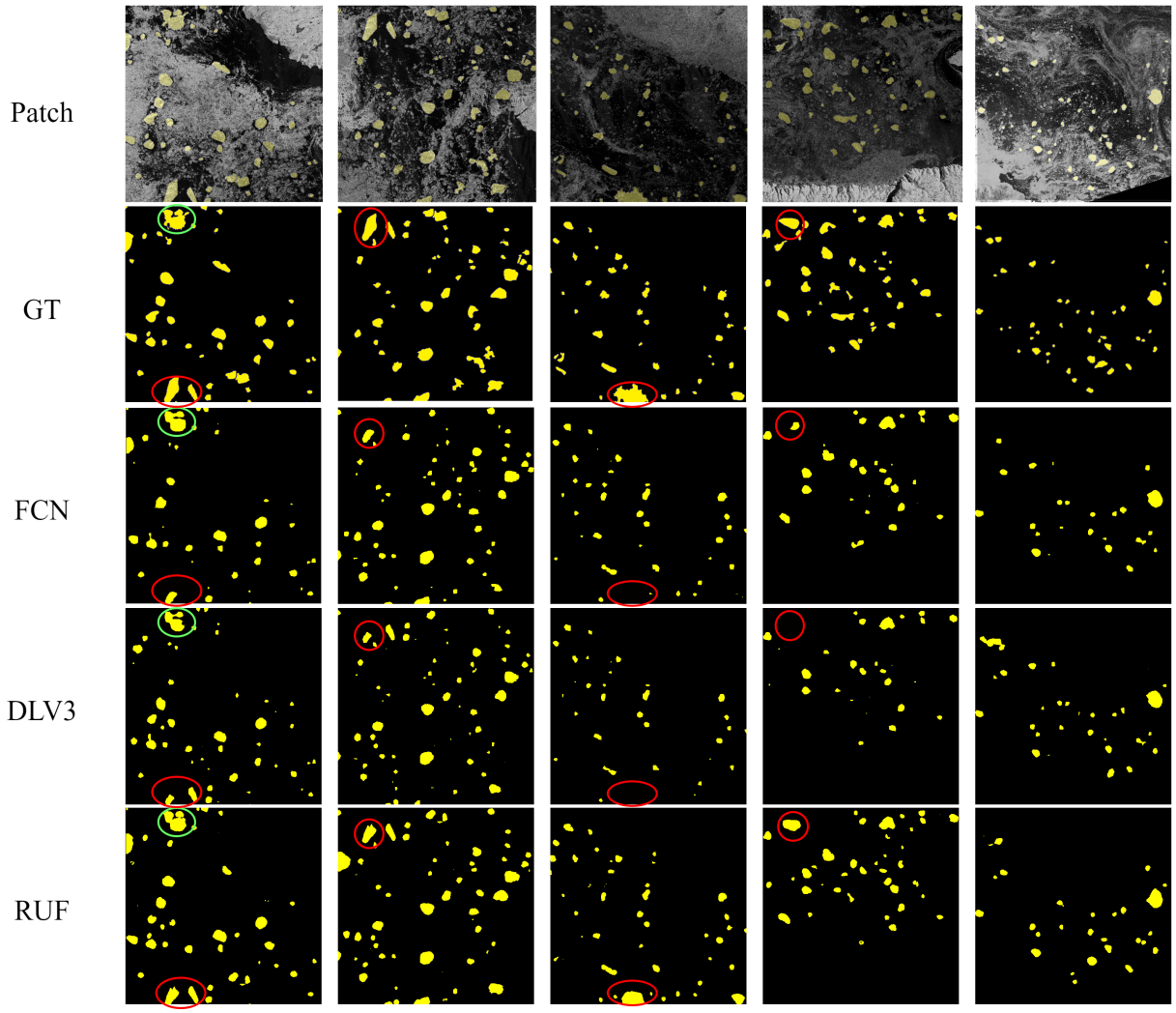


Figure 4.8: Comparison between the predictions of FCN-8, DeepLabV3, and RUF for the task of ice floe segmentation on patches from the test set. Red ovals highlight some segmentation results where both FCN and DLV3 are unable to properly segment a fully or partially visible ice floe. Green oval highlights a segmentation result where both FCN and DLV3 are unable to delineate floes in close proximity. It can be observed that the proposed RUF architecture is able to produce finer segmentation results as compared to the other frequently used segmentation methods in satellite imaging. Here, Patch and GT denotes original image patch input and ground truth respectively.

Chapter 5

Summary and Conclusion

Shipping vessel activity in the Arctic has observed steady growth both in the terms of volume and traffic over the past few years. According to the recent Arctic Climate Impact Assessment report, the extent and amount of sea ice in the Arctic region are decreasing gradually. The shipping industry has displayed a keen interest in the declining sea ice in the Canadian Arctic as the Northwest Passage of Canadian Arctic Archipelago (CAA) has the potential to become a seasonally viable shipping route. Although the Arctic is observing a decline in sea ice extent and thickness, the risks and hazards involved in sailing through the Arctic are still significant than in the past. One of the reasons for this behavior being that as the sea ice reduces, particularly in the Marginal Ice Zones (MIZs), the mobility and thus unpredictability of ice floes and ice-covered areas also increases in the coming future. Thus, MIZ detection is important in the northern seas to create safe and efficient navigable routes for ships. The Canadian Ice Service (CIS), is the leading authority which provides information regarding ice in Canada's navigable waters. CIS uses data from various data sources for the production of ice charts and the development of guidelines for mariners. One of the prominent data sources used for the task of sea ice monitoring is RADARSAT Constellation Mission equipped with Synthetic Aperture Radar (SAR) sensors. SAR provides high spatial resolution images irrespective of the daylight conditions due to the active nature of radars used for image acquisition. SAR images are not optical images and a certain amount of prerequisite knowledge is required to understand them. Sea ice experts are employed by CIS for interpretation of these images in order to produce the image analysis charts. To develop automated sea ice detection tools, previous studies have used various image processing, machine learning, and neural network approach as explained in Chapter 2. The main objective of this thesis is to develop automated methods for sea ice segmentation in SAR images. We presented two approaches

using convolutional neural networks (CNNs) to solve this task.

In the first approach, we developed a segmentation pipeline to detect MIZs in SAR images. We manually annotated a dataset of 160 images and extracted over 500 patches. Later, we explored the network-based transfer learning approach to train a CNN to classify the patches. The key point of our approach is the multi-scale patch technique. In this technique, we generate multiple segmentation masks for an image. Each individual segmentation mask is generated by feeding overlapping patches of a particular inference size to the network. Later, these segmentation masks are stacked together and thresholded to generate the final MIZ prediction mask. Results show that upon using a smaller inference patch size, the network generated predictions with a good recall but also generated false positive predictions, while a bigger inference patch size was able to classify negative samples easily. One of the reasons for this behavior may be due to the reason that upon using a bigger patch size, the network's decision is based on a larger spatial area. We observed that our method generates false predictions over severely wind roughened ocean water. To solve this issue we aim to use wind maps to mask the areas with high wind velocities. Wind maps have a very coarse spatial resolution, approximately 25 Km $\times$ 25 Km per pixel and further study is required to verify their applicability.

In the second approach, we developed a new segmentation architecture to detect individual ice floes in the MIZs of SAR images. We manually annotated a dataset of over 1800 floes. For an $1 \times H \times W$ input image, the network outputs an $1 \times H \times W$ segmentation mask. The network is based on a UNET with residual connections where a convolutional conditional random field (Conv-CRF) is used for post-processing. The whole network is trained end-to-end using a dual loss function, enabling both the UNET (primary model) and the Conv-CRF (secondary model) to co-optimize for the task. We observed that a vanilla UNET was able to some detect floes, but it generated many false-positive predictions and was unable to segment floes of bigger sizes. Upon adding residual connections, the model was able to detect bigger floes. Both the Conv-CRF joint training and dual loss functions steps helped the network to segment floes with finer boundaries. To gauge the performance of our architecture, we performed a quantitative comparison with various other popular segmentation models like FCN and DeepLabV3. RUF displayed an average improvement in performance with an F1 score of 74.52 (108.5% increment), mIOU of 66.26 (120.5% increment), and mPA of 77.88 (110.3% increment). In the future, we aim to experiment with attention networks to improve our results.

In this thesis, two approaches for automatic detection of MIZ and ice floes were explored and developed. The effectiveness of our models can be gauged from the obtained qualitative and quantitative results.

References

[1] Arctic open-water season grows. *Nature*, 527(7576):10–11, Nov 2015.

[2] Charu C Aggarwal et al. *Neural networks and deep learning*. Springer, 2018.

[3] Jonathan Andrews, David Babb, and David Barber. Climate change and sea ice: shipping in Hudson Bay, Hudson Strait, and Foxe Basin (1980–2016). *Elem Sci Anth*, 6(1), 2018.

[4] VV Asmus, EV Vasilenko, VV Zatyagalova, NP Ivanova, VA Krovotyntsev, AA Maksimov, and IS Trenina. Satellite Monitoring of Sea Ice Cover and Water Parameters for the Caspian Sea. *Russian Meteorology and Hydrology*, 43(10):686–696, 2018.

[5] Nicolas Audebert, Bertrand Le Saux, and Sébastien Lefèvre. Semantic segmentation of earth observation data using multimodal and multi-scale deep networks. In *asian conference on computer vision*, pages 180–196. Springer, 2016.

[6] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12):2481–2495, 2017.

[7] Jeffrey D Banfield and Adrian E Raftery. Ice floe identification in satellite images using mathematical morphology and clustering about principal curves. *Journal of the American Statistical Association*, 87(417):7–16, 1992.

[8] DG Barber, DG Babb, JK Ehn, W Chan, L Matthes, LA Dalman, Y Campbell, ML Harasyn, N Firoozy, N Theriault, et al. Increasing mobility of high Arctic sea ice increases marine hazards off the east coast of Newfoundland. *Geophysical Research Letters*, 45(5):2370–2379, 2018.

[9] Katherine R Barnhart, Christopher R Miller, Irina Overeem, and Jennifer E Kay. Mapping the future expansion of Arctic open water. *Nature Climate Change*, 6(3):280–285, 2016.

[10] Camilla Brekke, Benjamin Holt, Cathleen Jones, and Stine Skrunes. Discrimination of oil spills from newly formed sea ice by synthetic aperture radar. *Remote Sensing of Environment*, 145:1–14, 2014.

[11] C. Engler, R. Pelot, S. Okuribido-Malcolm, M. deHann, and O. Mussells, Proceedings of the Northern Marine Transportation Corridors Workshop, December 8, Vancouver, B.C., uO Research: Ottawa, ON. http://www.espg.ca/wp-content/uploads/2013/04/NMTC_Workshop_Proceedings_FINAL_REVISED.pdf, 2016. [Online; accessed 11-January-2020].

[12] Canadian Ice Service. Where sea ice is found. https://www.canada.ca/en/environment-climate-change/services/ice-forecasts-observations/latest-conditions/educational-resources/sea/where-sea-ice-is-found.html. Online; accessed 12-July-2020.

[13] Gongping Chen, Zhisheng Gao, Qiaolu Wang, and Qingqing Luo. U-net like deep autoencoders for deblurring atmospheric turbulence. *Journal of Electronic Imaging*, 28(5):053024, 2019.

[14] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4):834–848, 2017.

[15] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*, 2017.

[16] Angela Cheng, Barbara Casati, Adrienne Tivy, Tom Zagon, Jean-François Lemieux, and L Bruno Tremblay. Accuracy and inter-analyst agreement of visually estimated sea ice concentrations in Canadian Ice Service ice charts using single-polarization RADARSAT-2. *Cryosphere*, 14(4), 2020.

[17] R. Chénier, L. Abado, O. Sabourin, and Laurent Tardif. Northern marine transportation corridors: Creation and analysis of northern marine traffic routes in Canadian waters. *Transactions in GIS*, 21(6):1085–1097, 2017.

[18] David A Clausi and Bing Yue. Comparing cooccurrence probabilities and Markov random fields for texture analysis of SAR sea ice imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 42(1):215–228, 2004.

[19] J. C. Comiso. Abrupt decline in the Arctic winter sea ice cover. *Geophysical Research Letters*, 33(18), 2006.

[20] Colin L. V. Cooke and K. Andrea Scott. Estimating Sea Ice Concentration From SAR: Training Convolutional Neural Networks With Passive Microwave Data. *IEEE Transactions on Geoscience and Remote Sensing*, 57(7):4735–4747, 2019.

[21] RG Crane. Seasonal variations of sea ice extent in the Davis Strait-Labrador Sea area and relationships with synoptic-scale atmospheric circulation. *Arctic*, pages 434–447, 1978.

[22] G Crocker. Canadian ice service digital archive–regional charts: History accuracy and caveats. *Canadian Ice Service, Ottawa, ON, Canada, Tech. Rep*, 1, 2006.

[23] Claude Daley, Shadi Alawneh, Dennis Peters, Bruce Quinton, and Bruce Colbourne. Gpu modeling of ship operations in pack ice. In *International Conference and Exhibition on Performance of Ships and Structures in Ice, Banff Alberta, Canada September*, pages 20–23, 2012.

[24] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009.

[25] Wolfgang Dierking. Sea ice monitoring by synthetic aperture radar. *Oceanography*, 26(2):100–111, 2013.

[26] KF Drinkwater. Physical oceanography of Hudson Strait and Ungava Bay. In *Elsevier oceanography series*, volume 44, pages 237–264. Elsevier, 1986.

[27] Leif EB Eriksson, Karin Borenäs, Wolfgang Dierking, Anders Berg, Maurizio Santoro, Per Pemberton, Henrik Lindh, and Bengt Karlson. Evaluation of new spaceborne SAR sensors for sea-ice monitoring in the Baltic Sea. *Canadian Journal of Remote Sensing*, 36(sup1):S56–S73, 2010.

[28] G. Crocker. Canadian Ice Service Digital Archive-Regional Charts: History, Accuracy and Caveats, Ottawa, Ontario, 2006.

[29] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Region-based convolutional networks for accurate object detection and segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 38(1):142–158, 2015.

[30] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[31] Christian Haas et al. Dynamics versus thermodynamics: The sea ice thickness distribution. *Sea ice: An introduction to its physics, chemistry, biology and geology*, 1:82–111, 2003.

[32] Richard J Hall, Nick Hughes, and Peter Wadhams. A systematic method of obtaining ice concentration measurements from ship-based observations. *Cold Regions Science and Technology*, 34(2):97–102, 2002.

[33] Martti Hallikainen and Dale P Winebrenner. The physical basis for sea ice remote sensing. *GMS*, 68:29–46, 1992.

[34] Joakim Haugen, Lars Imsland, Sveinung Løset, Roger Skjetne, et al. Ice observer system for ice management operations. In *The Twenty-first International Offshore and Polar Engineering Conference*. International Society of Offshore and Polar Engineers, 2011.

[35] Donna Haverkamp, Leen Kiat Soh, and Costas Tsatsoulis. A comprehensive, automated approach to determining sea ice thickness from SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 33(1):46–57, 1995.

[36] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, pages 1026–1034, 2015.

[37] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2016.

[38] Hans-Martin Heyn, Mogens Blanke, and Roger Skjetne. Ice condition assessment using onboard accelerometers and statistical change detection. *IEEE Journal of Oceanic Engineering*, 2019.

[39] Hans-Martin Heyn, Martin Knoche, Qin Zhang, and Roger Skjetne. A system for automated vision-based sea-ice concentration detection and floe-size distribution indication from an icebreaker. In *International Conference on Offshore Mechanics and Arctic Engineering*, volume 57762, page V008T07A012. American Society of Mechanical Engineers, 2017.

[40] Marie Hoekstra, Mingzhe Jiang, David A Clausi, and Claude Duguay. Lake Ice-Water Classification of RADARSAT-2 Images by Integrating IRGS Segmentation with Pixel-Based Random Forest Labeling. *Remote Sensing*, 12(9):1425, 2020.

[41] Benjamin Holt and Seelye Martin. The effect of a storm on the 1992 summer sea ice cover of the Beaufort, Chukchi, and East Siberian Seas. *Journal of Geophysical Research: Oceans*, 106(C1):1017–1032, 2001.

[42] Christopher Houser and William A Gough. Variations in sea ice in the Hudson Strait: 1971-1999. *Polar Geography*, 27(1):1–14, 2003.

[43] Xiaodan Hu, Mohamed A Naiel, Alexander Wong, Mark Lamm, and Paul Fieguth. Runet: A Robust UNet Architecture for Image Super-Resolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 0–0, 2019.

[44] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[45] David H Hubel and Torsten N Wiesel. Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology*, 148(3):574, 1959.

[46] Byongjun Hwang, Jinchang Ren, Samuel McCormack, Craig Berry, Ismail Ben Ayed, Hans C Graber, and Erchan Aptoula. A practical algorithm for the retrieval of floe size distribution of arctic sea ice from high-resolution satellite Synthetic Aperture Radar imagery. *Elem Sci Anth*, 5, 2017.

[47] Forrest Iandola, Matt Moskewicz, Sergey Karayev, Ross Girshick, Trevor Darrell, and Kurt Keutzer. Densenet: Implementing efficient convnet descriptor pyramids. *arXiv preprint arXiv:1404.1869*, 2014.

[48] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.

[49] J. Dawson and L. Porta. Analysis of Marine Traffic along Canada's Coasts Phase 2 – Part 1: Factors Influencing Arctic Traffic. https://cradpdf.drdc-rddc.gc.ca/PDFS/unc125/p537704_A1b.pdf, 2013. [Online; accessed 11-January-2020].

[50] Kevin Jarrett, Koray Kavukcuoglu, Marc'Aurelio Ranzato, and Yann LeCun. What is the best multi-stage architecture for object recognition? In *2009 IEEE 12th International Conference on Computer Vision*, pages 2146–2153, 2009.

[51] A Malin Johansson, Camilla Brekke, Gunnar Spreen, and Jennifer A King. X-, C-, and L-band SAR signatures of newly formed sea ice in Arctic leads during winter and spring. *Remote Sensing of Environment*, 204:162–180, 2018.

[52] Juha Karvonen. Baltic sea ice concentration estimation using SENTINEL-1 SAR and AMSR2 microwave radiometer data. *IEEE Transactions on Geoscience and Remote Sensing*, 55(5):2871–2883, 2017.

[53] Juha Karvonen, Lijian Shi, Bin Cheng, Markku Similä, Marko Mäkynen, and Timo Vihma. Bohai sea ice parameter estimation based on thermodynamic ice model and Earth observation data. *Remote Sensing*, 9(3):234, 2017.

[54] Zoltan Kato, Marc Berthod, and Josiane Zerubia. A hierarchical Markov random field model and multitemperature annealing for parallel image classification. *Graphical models and image processing*, 58(1):18–37, 1996.

[55] A Keinonen and I Robbins. Icebreaker characteristics synthesis, icebreaker performance models, seakeeping, icebreaker escort. *Icebreaker Escort Model User's Guide: Report Prepared for Transport Development Centre Canada (TP12812E)*, 3:49, 1998.

[56] V Komarov, S Wang, and J Tang. Permittivity and measurements. *Encyclopedia of RF and microwave engineering*, 2005.

[57] Philipp Krähenbühl and Vladlen Koltun. Efficient inference in fully connected crfs with gaussian edge potentials. In *Advances in neural information processing systems*, pages 109–117, 2011.

[58] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[59] Philipp Krähenbühl and Vladlen Koltun. Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials, 2012.

[60] Ivana Kubat and Denise Sudom. Ship safety and performance in pressured ice zones: Captains' responses to questionnaire. *NRC Report CHC-TR-056*, 2008.

[61] G. Spreen L. Kaleschke and G.Heygster. Sea ice remote sensing using AMSR-E 89-GHz channels. *Journal of Geophysical Research*, 113:C02S03, January 2008.

[62] Yang Lei, Sibo Tian, Xiuxiu He, Tonghe Wang, Bo Wang, Pretesh Patel, Ashesh B Jani, Hui Mao, Walter J Curran, Tian Liu, et al. Ultrasound prostate segmentation based on multidirectional deeply supervised V-Net. *Medical physics*, 46(7):3194–3206, 2019.

[63] Steven Leigh, Zhijie Wang, and David A Clausi. Automated ice–water classification using dual polarization SAR satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(9):5529–5539, 2013.

[64] Yufeng Lin, Adolf KY Ng, and Mawuli Afenyo. Climate change, a double-edged sword: The case of Churchill on the Northwest Passage. In *Maritime Transport and Regional Sustainability*, pages 223–235. Elsevier, 2020.

[65] Jiange Liu, Katharine Andrea Scott, and Paul W Fieguth. Detection of marginal ice zone in Synthetic Aperture Radar imagery using curvelet-based features: a case study on the Canadian East Coast. *Journal of Applied Remote Sensing*, 13(1):014505, 2019.

[66] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.

[67] Peng Lu and Zhijun Li. A Method of Obtaining Ice Concentration and Floe Size From Shipboard Oblique Sea Ice Images. *IEEE Transactions on Geoscience and Remote Sensing*, 48(7):2771–2780, 2010.

[68] Nathanael Melia, Keith Haines, and Ed Hawkins. Sea ice decline and 21st century trans-Arctic shipping routes. *Geophysical Research Letters*, 43(18):9720–9728, 2016.

[69] MEPC. International code for ships operating in polar waters (polar code). http://www.imo.org/en/MediaCentre/HotTopics/polar/Documents/POLAR%20CODE%20TEXT%20AS%20ADOPTED.pdf, 2015. Online; accessed 12-July-2020.

[70] Ivan Metrikin, Sveinung Løset, Nils Albert Jenssen, and Sofien Kerkeni. Numerical simulation of dynamic positioning in ice. *Marine Technology Society Journal*, 47(2):14–30, 2013.

[71] Hrushikesh Mhaskar, Qianli Liao, and Tomaso Poggio. Learning functions: when is deep better than shallow. *arXiv preprint arXiv:1603.00988*, 2016.

[72] Arctic Monitoring and Assessment Program (AMAP). Snow, water, ice, and permafrost in the Arctic (SWIPA), 2017.

[73] Arctic Monitoring and Assessment Programme. Snow, water, ice and permafrost in the arctic (swipa): Climate change and the cryosphere, 2011.

[74] Vinod Nair and Geoffrey E. Hinton. Rectified Linear Units Improve Restricted Boltzmann Machines. In *Proceedings of the 27th International Conference on Machine Learning*, pages 807–814, 2010.

[75] Adolf KY Ng, Jonathan Andrews, David Babb, Yufeng Lin, and Austin Becker. Implications of climate change for shipping: Opening the Arctic seas. *Wiley Interdisciplinary Reviews: Climate Change*, 9(2):e507, 2018.

[76] Paul Nicklen. Spring meltwater pools form on top of the sea ice.

[77] Willy Ostreng, Karl Magnus Eger, Brit Fløistad, Arnfinn Jørgensen-Dahl, Lars Lothe, Morten Mejlænder-Larsen, and Tor Wergeland. *Shipping in Arctic waters: a comparison of the Northeast, Northwest and trans polar passages*. Springer Science & Business Media, 2013.

[78] Andrew Palmer. *Arctic offshore engineering*. World Scientific, 2013.

[79] Jeong-Won Park, Anton A Korosov, Mohamed Babiker, Joong-Sun Won, Morten W Hansen, and Hyun-Cheol Kim. Classification of Sea Ice Types in Sentinel-1 SAR images. *The Cryosphere Discussions*, pages 1–23, 2019.

[80] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala. Pytorch: An Imperative Style, High-Performance Deep Learning Library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. dAlché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.

[81] Paul R Pinet. *Invitation to oceanography*. Jones & Bartlett Learning, 2019.

[82] L. Pizzolato, S. E. L. Howell, C. Derksen, J. Dawson, and L. Copland. Changing sea ice conditions and marine transportation activity in Canadian Arctic waters between 1990 and 2012. *Climatic Change*, 123(2):161–173, 2014.

[83] P Rampal, Jérôme Weiss, and D Marsan. Positive trend in the mean speed and deformation rate of Arctic sea ice, 1979–2007. *Journal of Geophysical Research: Oceans*, 114(C5), 2009.

[84] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.

[85] Jinchang Ren, Byongjun Hwang, Paul Murray, Soumitra Sakhalkar, and Samuel McCormack. Effective SAR sea ice image segmentation and touch floe separation using a combined multi-stage approach. In *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 1040–1043. IEEE, 2015.

[86] Christopher Dahlin Rodin and Tor Arne Johansen. Accuracy of sea ice floe size observation from an aerial camera at slant angles. In *2017 Workshop on Research, Education and Development of Unmanned Aerial Systems (RED-UAS)*, pages 216–221. IEEE, 2017.

[87] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[88] Mohamed Ben Salah, Amar Mitiche, and Ismail Ben Ayed. Multiregion image segmentation by parametric kernel graph cuts. *IEEE Transactions on Image Processing*, 20(2):545–557, 2010.

[89] Stein Sandven, Ola M Johannessen, and Kjell Kloster. Sea ice monitoring by remote sensing. *Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation*, 2006.

[90] FJ Saucier, S Senneville, S Prinsenberg, F Roy, G Smith, P Gachon, D Caya, and R Laprise. Modelling the sea ice-ocean seasonal cycle in Hudson Bay, Foxe Basin and Hudson Strait, Canada. *Climate Dynamics*, 23(3-4):303–326, 2004.

[91] Mohamed Sayed, Anne Barker, et al. Numerical simulations of ice interaction with a moored structure. In *OTC Arctic Technology Conference*. Offshore Technology Conference, 2011.

[92] Mohamed Sayed, Ivana Kubat Kubat, Brian Wright, et al. Numerical simulations of ice forces on the kulluk: The role of ice confinement, ice pressure and ice management. In *OTC Arctic Technology Conference*. Offshore Technology Conference, 2012.

[93] Bernd Scheuchl, Dean Flett, Ron Caves, and Ian Cumming. Potential of RADARSAT-2 data for operational sea ice monitoring. *Canadian Journal of Remote Sensing*, 30(3):448–461, 2004.

[94] Boris Sekachev, Nikita Manovich, zhiltsov max, Andrey Zhavoronkov, Dmitry Kalinin, Ben Hoff, TOsmanov, Artyom Zankevich, DmitriySidnev, Maksim Markelov, Johannes222, Mathis Chenuet, a andre, telenachos, Aleksandr Melnikov, Dmitry Kruchinin, Jijoong Kim, Liron Ilouz, Nikita Glazov, Priya4607, Rush Tehrani, Seungwon Jeong, Vladimir Skubriev, Sebastian Yonekura, vugia truong, zliang7, lizhming, and Tritin Truong. opencv/cvat: v1.1.0-beta, August 2020.

[95] Canadian Hydrographic Service. *Atlas des courants*. Canadian Hydrographic Service= Service hydrographique du Canada, 1983.

[96] G. E. Shephard, K. Dalen, R. Peldszus, S. Aparício, L. Beumer, R. Birkeland, N. Gkikas, M. Kourantidou, P. Ktenas, P. W. Linde, et al. Assessing the added value of the recent declaration on unregulated fishing for sustainable governance of the central Arctic Ocean. *Marine Policy*, 66:50–57, 2016.

[97] Lijian Shi, Juha Karvonen, Bin Cheng, Timo Vihma, Mingsen Lin, Yu Liu, Qimao Wang, and Yongjun Jia. Sea ice thickness retrieval from SAR imagery over bohai sea. In *2014 IEEE Geoscience and Remote Sensing Symposium*, pages 4864–4867. IEEE, 2014.

[98] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[99] VG Smirnov and IA Bychkova. Satellite monitoring of ice features to ensure safety of offshore operations in the Arctic seas. *Izvestiya, Atmospheric and Oceanic Physics*, 51(9):935–942, 2015.

[100] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958, 2014.

[101] Adam Steer, Anthony Worby, and Petra Heil. Observed changes in sea-ice floe size distribution during early summer in the western Weddell Sea. *Deep Sea Research Part II: Topical Studies in Oceanography*, 55(8-9):933–942, 2008.

[102] Fiammetta Straneo and François Saucier. The outflow from Hudson Strait and its contribution to the Labrador Current. *Deep Sea Research Part I: Oceanographic Research Papers*, 55(8):926–946, 2008.

[103] Fiammetta Straneo and François J Saucier. The arctic–subarctic exchange through Hudson Strait. In *Arctic–subarctic ocean fluxes*, pages 249–261. Springer, 2008.

[104] C. Strong. Atmospheric influence on Arctic marginal ice zone position and width in the Atlantic sector, February–April 1979–2010. *Climate dynamics*, 39(12):3091–3102, 2012.

[105] Carole H Sudre, Wenqi Li, Tom Vercauteren, Sebastien Ourselin, and M Jorge Cardoso. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep learning in medical image analysis and multimodal learning for clinical decision support*, pages 240–248. Springer, 2017.

[106] Marvin T. T. Teichmann and Roberto Cipolla. Convolutional CRFs for Semantic Segmentation, 2018.

[107] David N Thomas. *Sea ice*. John Wiley & Sons, 2017.

[108] GW Timco, K Croasdale, and B Wright. An overview of first-year sea ice ridges. *PERD/CHC report*, pages 5–112, 2000.

[109] Takenobu Toyota, Christian Haas, and Takeshi Tamura. Size distribution and shape properties of relatively small sea-ice floes in the Antarctic marginal ice zone in late winter. *Deep Sea Research Part II: Topical Studies in Oceanography*, 58(9-10):1182–1193, 2011.

[110] Costas Tsatsoulis and Ronald Kwok. *Analysis of SAR data of the polar oceans: recent advances*. Springer Science & Business Media, 2012.

[111] Fawwaz T Ulaby, Richard K Moore, and Adrian K Fung. Microwave remote sensing: Active and passive. volume 1-microwave remote sensing fundamentals and radiometry. 1981.

[112] Peter Wadhams. How does Arctic sea ice form and decay. *Retrieved May*, 19(2008):275–332, 2003.

[113] Lei Wang, K Andrea Scott, Linlin Xu, and David A Clausi. Sea ice concentration estimation during melt from dual-pol SAR scenes using deep convolutional neural networks: A case study. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8):4524–4533, 2016.

[114] Panqu Wang, Pengfei Chen, Ye Yuan, Ding Liu, Zehua Huang, Xiaodi Hou, and Garrison Cottrell. Understanding convolution for semantic segmentation. In *2018 IEEE winter conference on applications of computer vision (WACV)*, pages 1451–1460. IEEE, 2018.

[115] Qingkai Wang, Zhijun Li, Peng Lu, Ruibo Lei, and Bin Cheng. 2014 summer Arctic sea ice thickness and concentration from shipborne observations. *International Journal of Digital Earth*, 12(8):931–947, 2019.

[116] B Wright et al. Evaluation of full scale data for moored vessel stationkeeping in pack ice. *PERD/CHC Report*, (26-200), 1999.

[117] Tao Xie, William Perrie, Caiying Wei, and Li Zhao. Discrimination of open water from sea ice in the labrador sea using quad-polarized synthetic aperture radar. *Remote Sensing of Environment*, 247:111948, 2020.

[118] Linlin Xu, Mohammad Javad Shafiee, Alexander Wong, and David A Clausi. Fully connected continuous conditional random field with stochastic cliques for dark-spot detection in SAR imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(7):2882–2890, 2016.

[119] Ma Yi-de, Liu Qing, and Qian Zhi-Bai. Automated image segmentation using improved PCNN model based on cross-entropy. In *Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, 2004.*, pages 743–746. IEEE, 2004.

[120] Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. How transferable are features in deep neural networks? In *Advances in neural information processing systems*, pages 3320–3328, 2014.

[121] Sergey Zagoruyko and Nikos Komodakis. Wide residual networks. *arXiv preprint arXiv:1605.07146*, 2016.

[122] Natalia Yu Zakhvatkina, Vitaly Yu Alexandrov, Ola M Johannessen, Stein Sandven, and Ivan Ye Frolov. Classification of sea ice types in ENVISAT synthetic aperture

radar images. *IEEE Transactions on Geoscience and Remote Sensing*, 51(5):2587–2600, 2012.

[123] Xiuwei Zhang, Jiaojiao Jin, Zeze Lan, Chunjiang Li, Minhao Fan, Yafei Wang, Xin Yu, and Yanning Zhang. Icenet: A Semantic Segmentation Deep Network for River Ice by Fusing Positional and Channel-Wise Attentive Features. *Remote Sensing*, 12(2):221, 2020.

[124] Zhengxin Zhang, Qingjie Liu, and Yunhong Wang. Road Extraction by Deep Residual U-Net. *IEEE Geoscience and Remote Sensing Letters*, 15(5):749–753, May 2018.

[125] Shuai Zheng, Sadeep Jayasumana, Bernardino Romera-Paredes, Vibhav Vineet, Zhizhong Su, Dalong Du, Chang Huang, and Philip HS Torr. Conditional random fields as recurrent neural networks. In *Proceedings of the IEEE international conference on computer vision*, pages 1529–1537, 2015.

[126] Yi-Tong Zhou and Rama Chellappa. Computation of optical flow using a neural network. In *ICNN*, pages 71–78, 1988.

[127] Tingting Zhu, Fei Li, Georg Heygster, and Shengkai Zhang. Antarctic Sea-Ice Classification Based on Conditional Random Fields From RADARSAT-2 Dual-Polarization Satellite Images. *IEEE Journal of Selected topics in Applied Earth observations and Remote Sensing*, 9(6):2451–2467, 2016.

[128] É. Laurent, P. Ronald and E. Cecilia. Analysis of Marine Traffic along Canada's Coasts Phase 2 – Part 2: A spatio-temporal simulation model for forecasting marine traffic in the Canadian Arctic in 2020. https://cradpdf.drdc-rddc.gc.ca/PDFS/unc137/p538451_A1b.pdf, 2013. [Online; accessed 11-January-2020].